GENE EXPRESSION IN THE HIPPOCAMPUS: IDENTIFICATION OF A NOVEL CYTOCHROME P-450.

GENEVIEVE STAPLETON

Thesis presented for the degree of Doctor of Philosophy

University of Edinburgh

1994

I declare that this work is my own, except where otherwise stated.

Genevieve Stapleton

To my parents

ACKNOWLEDGEMENTS

First of all, I thank Professor Rick Lathe for providing me with the opportunity to come to Edinburgh (let's not mention the weather) and the Centre for Genome Research.

My supervision over the last four years has been a "mixed bag", and I thank Patrizia Lavia, John Mason and Rick Lathe for their individual contributions.

I was very fortunate to find myself in the same lab as Melville Richardson. Without her, I'd still be learning mini-preps, but more importantly perhaps, she taught me that "licorice is good for concentration". I was lucky also to share the lab with John Mason, Ben Pickard and Muriel Steel, and I am indebted to them for their invaluable help, contributions, and support.
Overall, however, the uniting feature of the "hippo lab" was our dedication to the appreciation of ice-cream (and anything sweet really). We were thorough!

Thank-you also to Professor Richard Morris and colleagues for their collaboration and input to the project.

In addition to lab members, I thank Caroline Herron and Lesley Forrester for helpful comments and proof-reading.

Thanks also to Patrizia Lavia and Patrizia Somma, Anne O'Garra and Hope Rugo (for the initial push), friends and family in Melbourne (especially Manuela and Kate), Tom Gardner, Otto Windl and Branko Zevnik (for "5 o'clock drinks") and Pat Hope.

Finally a big thank-you to Christopher McInerny ... for just about everything.

ABBREVIATIONS

The recommended nomenclature (Nelson *et al*, (1993)) for cytochrome P-450 genes and their products has been adopted throughout this work.

Standard abbreviations for nucleotides, and both the single letter and three-letter abbreviations for amino acids were used throughout the text.

Standard abbreviations and symbols recommended by the IUPAC-IUB Commission on Biochemical Nomenclature have generally been used.

Non-standard abbreviations are listed below.

APS	ammonium persulphate
ATP	adenosine triphosphate
bp	base pair(s)
cDNA	complementary DNA
dNTP	deoxynucleotide triphosphate
dH_2O	distilled water
DEPC	diethylpyrocarbonate
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
DTT	dithiothrietol (Cleland's reagent)
EDTA	ethylenediaminotetraacetic acid (disodium salt)
EGTA	ethylene glycol-bis(b-aminoethyl ether) N,N,N',N'-tetraacetic acid
ES	embryonic stem
HEPES	N-2-hydroxyethylpiperazine-N-2-ethanesulphonic acid
IPTG	isopropyl β-D-thiogalactopyranoside
kb	kilobase
kD	kilodalton
Klenow	DNA polymerase I, large fragment
MOPS	3-[N-Morpholino] propanesulphonic acid
mRNA	messenger ribonucleic acid

nt	nucleotide(s)
O.D.	optical density
oligonucleotide	oligodeoxyribonucleotide
PAGE	polyacrylamide gel electrophoresis
PEG	polyethylene glycol
RNA	ribonucleic acid
RNaseA	ribonuclease A
rpm	revolutions per minute
SDS	sodium dodecyl sulphate
SM	phage resuspension buffer
TAE	Tris.Cl-acetate electrophoresis buffer
TBE	Tris.Cl-borate electrophoresis buffer
TE	10 mM Tris.Cl (pH 7.4); 1 mM EDTA
TEMED	N,N,N',N'-tetramethyl-ethylenediamine
Tris	tris (hydroxymethyl) aminomethane
X-gal	5-bromo-4-chloro-3-indolyl-B-D-galactoside

CONTENTS

Appendix I

A summary of known cytochrome P-450 expression or activity in brain.

Appendix II

Regulation of housekeeping gene expression: "Cell type-specific interactions of transcription factors with a housekeeping promoter *in vivo*".

ABSTRACT

Neuropsychological studies in amnesiac patients and surgical lesioning in animals have demonstrated that the integrity of the hippocampal formation is essential for the establishment of some but not all types of memory. The exact mechanism by which the hippocampus contributes to memory formation and the underlying molecular and biochemical processes are not understood. One approach to understanding hippocampal function is through transgenesis. The identification of genes specifically expressed in this formation would provide promoters capable of directing reporter gene expression to the hippocampus. Conditional cellular ablation of the hippocampus, or subregions thereof, could address more precisely its role in memory. In addition, molecular mechanisms of memory and hippocampal function may be investigated through the targeted expression of suitably modified components of signalling pathways thought to be involved.

This thesis aims to identify genes expressed in the hippocampus. Such genes would permit targeting hippocampus-specific gene expression, while their identity and function within the hippocampus may also be characterised.

Differential hybridisation techniques, employing cDNA libraries prepared from hippocampus or other brain regions, have been used in an attempt to identify genes expressed either specifically or predominantly in the hippocampus. One cDNA clone isolated, Hct1 (hippocampal transcript), is further characterised. The expression pattern of Hct1, by Northern analysis and *in situ* hybridisation, shows significantly enriched expression within the adult rat hippocampus, as well as weak expression in the liver. However, Hct1 expression in the adult mouse brain is not hippocampus enriched and is found throughout the central nervous system. Although there is no clear justification of using the Hct1 regulatory elements in transgenic mouse experiments, there may be possibilities in rat. Sequencing of cDNA and genomic clones of Hct1 identified the encoded protein as a novel member of the cytochrome P-450 superfamily, a large group of enzymes involved in the oxidative metabolism of a variety of steroids, fatty acids and xenobiotics. Hct1 may represent a new family of cytochrome P-450 genes. The possible role of this enzyme in hippocampal function is discussed.

CHAPTER ONE

INTRODUCTION

1.1 Memory and the hippocampus

1.1.1 Memory

Learning, perhaps the major means by which behaviour is determined by the environment, can be defined as the acquisition of knowledge from the environment, while memory is the ability to store and subsequently recall such information. However such simple definitions do not distinguish between the different types of memory that can be recognised (Figure 1.1). For example, depending on the duration of the change in a behaviour, a memory may be termed short term (seconds to hours) or long term (days to years). Different types of memory have become apparent through a number of different paradigms (reviewed in Squire, 1992). Less complex invertebrate models have been used to demonstrate simple forms of learning, while more complex types of memories have been defined through the analysis of human amnesia and animal models. Moreover, these latter studies have shown that the formation of memory is not confined to a single centre within the brain, rather, that different types of memories are effected by different regions of the brain.

Two defensive behaviours in the marine snail *Aplysia californica* have been used to investigate certain types of learning (reviewed in Kandel, 1991). One behaviour is the siphon-gill withdrawal reflex where a tactile or electrical stimulus to the siphon results in reflex withdrawal of the siphon and gill into the mantle cavity. The tail-siphon reflex, where the stimulus is delivered to the tail, similarly results in the coordinated withdrawal of the tail, siphon and gill. It should however be noted that, while *Aplysia* is the example used here,

different forms of learning can be demonstrated in a variety of organisms, including flies (reviewed in Davis, 1993) and mammals.

Two simple forms of memory which can be demonstrated with these behavioural reflexes are habituation and sensitisation (Byrne *et al.*, 1991). Habituation is the learned, reduced response from a subject to a stimulus which is not noxious. Thus, a mild, tactile stimulus to the tail of *Aplysia* will result in habituation of the response after repeated stimulation. The subject learns that the stimulus is not a threat and reduces the reflex response on repeated stimulation. Sensitisation refers to an enhancement in the behavioural response to a strong, noxious stimulus. Sensitisation in the tail-siphon reflex would be measured as an increase in the duration of withdrawal of the siphon, that is, the animal responds more effectively to a previously encountered noxious stimulus. Both sensitisation and habituation are examples of non-associative, reflexive learning. Non-associative learning results when the organism is exposed once or repeatedly to one stimulus only, occurring without any specific temporal relationship between a test stimulus and a second reinforcing stimulus.

Associative learning requires the organism to learn the relationship of one stimulus to another (Byrne *et al.*, 1990). Classical conditioning is an example of associative learning where the subject learns the temporal association of two distinct stimuli. The gill-withdrawal reflex of *Aplysia* can be classically conditioned by pairing a conditioned stimulus with an unconditioned stimulus to the tail. When repeatedly paired, the gill withdrawal response becomes much greater than if either stimulus is delivered alone, and is retained for several days (Byrne *et al.*, 1990).

While classical conditioning establishes a contingency between a stimulus and a reinforcement, in contrast, operant conditioning produces an association between a response of the subject and reinforcement. The head-waving behaviour of *Aplysia* has been used by Cook and Carew (1986) to demonstrate operant conditioning. Head-waving is a naturally-occurring behaviour used by *Aplysia* to explore their environment. In this example, a strong light is used as a noxious stimulus, and is shone whenever the snail turns its head in a particular direction, while moving its head in the other direction will

result in no aversive signal. *Aplysia* can be operantly trained to modify the direction of their head-waving behaviour in order to avoid the noxious stimulus. Thus operant conditioning relies on the behaviour of the subject in response to a stimulus.

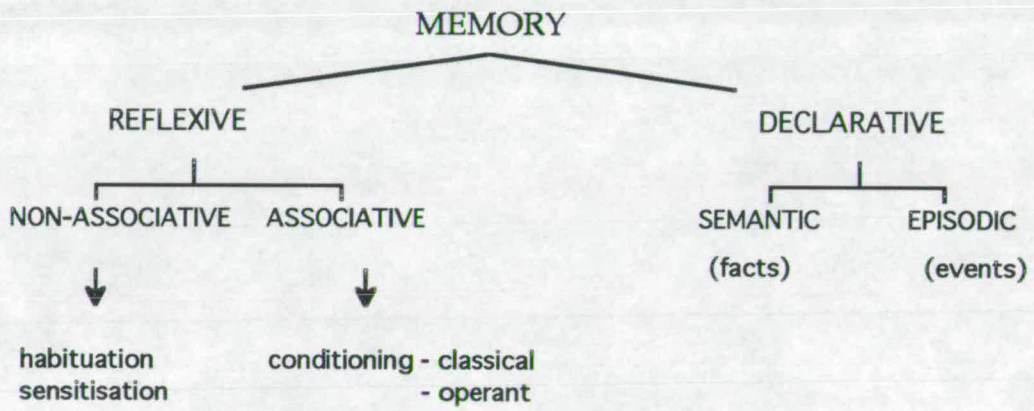


Figure 1.1 A classification of different types of memory. The classification presented is not exhaustive, rather it is intended to illustrate the complexity of memory, and some of the types which can be distinguished experimentally.

The types of learning so far described can be classified according to the experimental procedure used to establish them. However, all of these can be grouped together on the basis of the type of knowledge acquired and its formation in each case. As shown in Figure 1.1, habituation, sensitisation, and conditioning are all reflexive, or non-declarative, types of memory. These memories are formed through repetition over many trials, and are automatic, motor responses of a subject to avoid a noxious stimulus or to encourage a pleasant one. Such memories require no awareness or cognitive understanding.

In contrast, declarative memory includes the learning of facts, events and experiences, and the ability to report or 'declare' such memories. Such declarative memories include the ability to

remember faces and spatial layouts, where the 'declaration' is not necessarily verbal.

1.1.2 Memory in vertebrates

Experimental systems such as *Aplysia* are to the advantage of neuroscientists as their relatively simple nervous systems afford examination of the neurons and circuitry involved in a particular type of learning. More complex types of declarative memory can be defined in the nervous systems of vertebrates where the size and complexity of the nervous system is greatly increased. Thus, studies on amnesiac patients, and surgical manipulation of the nervous system of rodents and primates, have been used to address higher order cognition. In addition, certain forms of neurological dysfunction, where cognitive function is impaired, have been exploited in the study of processes underlying learning and memory in mammals.

(a) Human amnesia

Amnesia is defined as the impaired ability to learn new material (anterograde amnesia) against a background of intact intellectual ability. Retrograde amnesia is the loss of access to material stored before the onset of the memory impairment. As severely amnesiac patients can be entirely normal at some kinds of learning and memory, patients with a relatively selective memory deficit are of particular value and further define multiple memory systems. Thus, not only can the region of the brain responsible for the impairment be identified, particular types of memory formation can be assigned to that area.

In 1957, a severe but specific memory impairment was described in patient H.M. following bilateral surgical removal of the medial temporal lobe to alleviate severe epilepsy (Scoville and Milner, 1957). This lesion, known as H^+A^+, included the hippocampal formation (dentate gyrus, CA fields of the hippocampus proper, and subiculum), the amygdaloid complex, and adjacent cortical areas (entorhinal, perirhinal, periamygdaloid, and parahippocampal cortices). Unexpectedly, the surgery resulted in a selective impairment to

memory alone, leaving non-mnemonic functions intact. Long-term anterograde amnesia is severe in H.M., while immediate or "working" memory is spared. Retrograde amnesia is also restricted so that remote childhood memories are intact, but those more recent to the surgery date are lost. More specifically, however, the learning deficit in H.M. is in declarative types of memory. Procedural memory, where a skill is memorised through repetition (like learning to ride a bicycle) is non-declarative and is not affected in patients like H.M. who have hippocampal lesions (described in Cohen and Eichenbaum, 1993). While this large lesion locates the medial temporal lobe as a region of the brain involved in a specific cognitive function, it was not obvious which of the lost components (if not all) were essential for the lost function.

Patient R.B. provided evidence for a role of the hippocampus itself in memory, when global ischemia resulted in bilateral lesions of the full rostral-caudal extent of the CA1 region of the hippocampus (described in Squire, 1992). This very specific lesion resulted in a moderately severe memory impairment, with no other obvious intellectual deficits. Thus the hippocampus was shown to be a critical component of the memory system present in the medial temporal lobe of the brain. However, because R.B. was much less amnesiac than H.M., removal of other parts of the hippocampal formation may contribute to the severity of H.M.'s impairment. Thus, the CA2 and CA3 fields of the hippocampus, the dentate gyrus and the amygdala, as well as associated cortical regions may well also participate in memory processes.

(b) Animal models of learning and memory

Animal models, particularly the rat and primates, allow for the systematic investigation of anatomical structures important for memory. Primates have provided the best animal model to approximate human amnesia. This is mainly because several neuropsychological tests used for humans can be applied directly or adapted to primates. In addition, the life span of primates allows long

term memory and retrograde amnesia to be tested several years after surgery.

Initial experiments, intended to reproduce the lesion incurred by patient H.M., the H^+A^+ lesion, were followed by a series of more specific lesions in an attempt to define the critical regions within the medial temporal lobe (Table 1.1) (reviewed in Squire, 1991).

LESION	AREA REMOVED	SEVERITY OF DEFICIT
H^+A^+	hippocampus, amygdala and adjacent cortical areas*	+++
H^+	H^+A^+, leaving the amygdala, adjacent cortex and entorhinal cortex intact	++
A	amygdala	-
H^{++}	H^+ + anterior entorhinal, perirhinal cortex	+++
PRPH	perirhinal and parahippocampal cortex	+++

Table 1.1 A summary of surgical lesions in primates designed to analyse memory impairment. H refers to the hippocampus, A to the amygdala and $^+$ to the cortical regions adjacent to the hippocampus and amygdala. Memory impairment was determined using the 'delayed non-matching to sample' task. Recognition memory is tested by presenting an object to the subject, then after a delay the subject is presented with the same object and a novel one. The subject must pick the novel object to be rewarded (summarised from Squire, 1991).
* "adjacent cortex" includes periamygdaloid, perirhinal, entorhinal and parahippocampal cortices.

H^+A^+ lesions produce severe memory impairment in monkeys, but do not impair the acquisition or retention of motor skills. The H^+ lesion, which leaves the amygdala, adjacent cortex, and the entorhinal cortex unaffected, results in less severe memory impairment. This confirmed the role of the hippocampus but also confirmed a role for either the amygdala, associated cortical regions or both in memory formation. The A lesion and the H^{++} lesion established that disruption to the entorhinal cortex, the major input pathway into the hippocampus from the neocortex, also contributed to memory impairment. The A lesion selectively removed the amygdala without any resulting disruption to declarative memory. The H^{++} lesion extended the H^+ lesion (hippocampal formation and parahippocampal cortex) to include anterior entorhinal cortex and perirhinal cortex. The H^{++} lesion resulted in memory impairment equal in severity to the H^+A^+ lesion. Finally, the PRPH lesion, where only the perirhinal and parahippocampal cortex was removed, resulted in deficits similar to those of the H^+A^+ lesion. The PRPH effectively cut off all neocortical input into the entorhinal cortex and thus the hippocampus.

The lesion studies described confirmed a role for the hippocampus in the declarative memory system and identified the entorhinal cortex and its cortical input from the perirhinal and parahippocampal cortices as additional anatomical structures required for normal memory function.

(c) Neurodysfunction resulting in cognitive impairment

Memory impairment is associated with certain neuropathological conditions, the aetiology of which has also contributed to our current understanding of cognitive function. Alzheimer's disease is the most common form of dementia, resulting from chronic neuronal degeneration (Masters *et al.*, 1994). The pathology of Alzheimer's disease is associated with the deposition of amyloid plaques and the formation of neurofibrillary tangles, resulting in neuronal loss. In addition, memory impairment is an early and prominent

manifestation of the disease. A study of the medial temporal lobe in patients with Alzheimer's disease (Hyman *et al.*, 1984) revealed a specific cellular pattern of the pathology consistent with memory impairment. Cell types that are preferentially affected include those which provide the connections between the hippocampus and other regions of the brain. Cortical input into the hippocampus is from the projection neurons of the entorhinal cortex, while hippocampal output to the cortex is from projection neurons of the subiculum (reviewed in Amaral, 1993). Cell types invariantly affected in Alzheimer's disease included those of the subiculum and layers II, III and IV of the entorhinal cortex (Van-Hoesen *et al.*, 1991).

The destruction of the major input and output neurons of the hippocampus effectively isolates the structure from the rest of the brain. The disruption of memory processes in Alzheimer's patients further supports a role for the hippocampus as an integral part of memory formation. It should be noted, however, that Alzheimer's pathology is not exclusive to the above areas; other types of memory deficits may result from other damaged areas of the brain. In addition, the subiculum projects from the hippocampus to a number of targets including the amygdala and the cerebral cortex. The consequences of lack of innervation to these targets has not yet been addressed.

1.1.3 The hippocampus

Neuropsychological studies on amnesia models have shown that damage to the hippocampal formation produces an impairment of the declarative memory system. The hippocampus, together with the amygdala and the medial parts of the frontal, parietal and temporal cortical lobes, constitutes the limbic system, which is associated with learning, memory and emotion (described in Kandel, 1991).

To clarify terminology, it should be noted that hippocampus proper comprises the CA fields and the dentate gyrus, while the hippocampal formation includes the hippocampus, the entorhinal cortex and the subiculum.

The hippocampus, and thus hippocampus-dependent memory systems, has received further attention due to the relatively simple

structure of the formation. The hippocampus is laminar, allowing slice preparations for electrophysiology which conserve the circuitry of the structure. Electrophysiology and morphology define three major excitatory neuronal cell types in the hippocampus which connect afferent input from other parts of the brain (Figure 1.2). The major afferent input to the hippocampus is from the entorhinal cortex, although the septum and locus ceruleus also project to the hippocampus. The entorhinal cortex receives afferents from several neocortical regions and sends major projections from layers II and III through the perforant path to the hippocampus. Projections from layer II of the entorhinal cortex extend to the granule cells of the dentate gyrus and CA3 pyramidal cells, while projections from layer III innervate the CA1 field. Mossy fibre axons from the dentate granule cells project to CA3 cells, while CA3 pyramidal cells project via Schaffer collateral axons to area CA1. CA1 axons then project out of the hippocampus to the entorhinal cortex and the subiculum (summarised in Amaral, 1993).

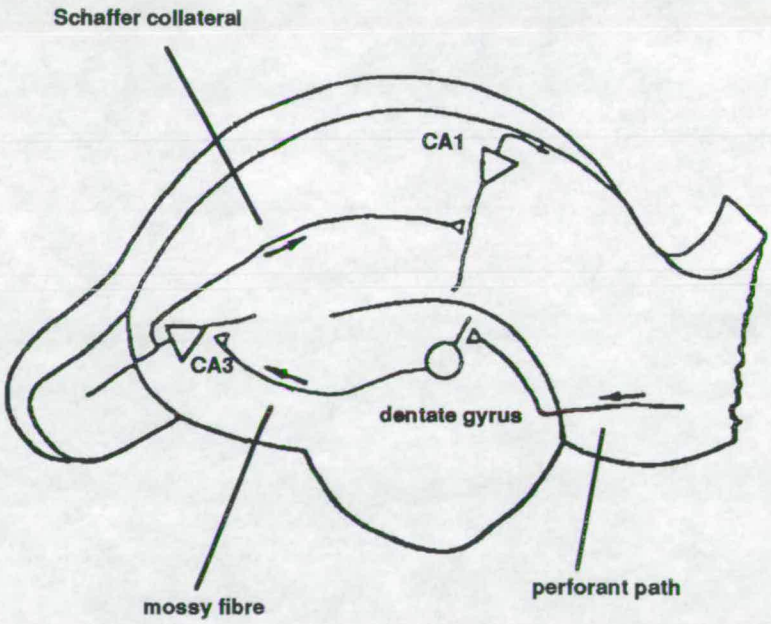


Figure 1.2 A diagram of the hippocampus, indicating the major excitatory neuronal subtypes and basic trisynaptic circuitry. Adapted from Bliss *et al.* (1990).

Further distinction of the cell types within the hippocampus requires the identification of genetic markers for specific cell types. GABAergic inhibitory interneurons within the hippocampus can be distinguished by the specific expression of GABA (γ-amino-butyric acid), however expression patterns of the calcium-binding proteins parvalbumin, calbindin D28k and cholecystokinin have identified subpopulations of GABAergic neurons in the hippocampus (Gulyas *et al.*, 1991). While the pyramidal cells of CA1-3 are morphologically indistinguishable, a serotonin receptor subtype described by Wisden *et al.* (1993) and Erlander *et al.* (1993) is expressed only in CA1 pyramidal cells within the hippocampus, distinguishing these cells from the rest of the structure. Dentate gyrus-specific gene expression has been reported following induction of synaptic activity (Nedivi *et al.*, 1993), further defining this cell type. Also, the "flip" and "flop" alternatively spliced mRNAs for the AMPA-responsive glutamate receptor subunit GluR-B are expressed differentially in CA pyramidal cells; GluR-B flip is expressed in CA3 cells and GluR-B flop is expressed in CA1 and dentate gyrus (Sommer *et al.*, 1990).

1.2 Long term potentiation

The cellular and molecular events underlying learning and memory are not understood. However, synapses have been shown to have adaptive qualities, for example, during the development of visual cortex synapses (reviewed in Shatz, 1990). Synaptic plasticity refers to structural and functional alterations of synaptic efficacy as a result of neuronal activity and is the underlying basis for the currently proposed model of learning.

Long-term potentiation (LTP) is an enduring increase in synaptic efficacy, first demonstrated by Bliss and Lømo (1973) in the

anaesthetised rabbit, then subsequently shown *in vitro* in hippocampus slice preparations (reviewed in Bliss and Collingridge, 1993). LTP is not exclusive to the hippocampus and has been demonstrated in a number of other brain regions including the visual cortex (Artola and Singer, 1987). LTP can be induced in a number of ways, though most typically by delivering a tetanus (a train of 50-100 stimuli at 100 Hz or more) to the presynaptic cell. The strength of synaptic transmission is measured by recording the excitatory postsynaptic potential (epsp); LTP is expressed as the increase in that postsynaptic response to constant test stimuli.

The major synaptic pathways in the hippocampus, the perforant path/dentate gyrus, mossy fibre/CA3, and Schaffer collateral/CA1, are all capable of exhibiting LTP. However the mechanisms underlying LTP in each are not necessarily the same, perhaps reflecting qualitatively different forms of synaptic plasticity. Of particular note is LTP produced at the mossy fibre/CA3 synapse, where, though these neurons are glutamatergic, neither NMDA receptor antagonists or calcium chelators injected into the postsynaptic cell appear to affect the induction of LTP (Zalutsky and Nicoll, 1990) (see section 1.2.1(a)).

1.2.1 Biochemistry of LTP

In the following section, discussion of LTP is restricted to the mechanisms thought to operate in CA1 of the hippocampus, or more specifically, at CA1 synapses receiving inputs from CA3 axons.

(a) The induction of LTP

The biochemical events which occur during the induction of CA1 LTP are initiated by an influx of calcium ions into the postsynaptic cell. Calcium entry has been shown to be critical as LTP can be prevented by injecting the Ca^{2+} chelator EGTA into postsynaptic neurons (Lynch *et al.*, 1983). Calcium ions enter through the NMDA (N-methyl-D-aspartate) receptor-associated channel. The NMDA channel is blocked by magnesium in a voltage-dependent manner. Fast synaptic transmission in the hippocampus is mediated by glutamate activation

of AMPA (alpha-amino-3-hydroxy-5-methyl-4 isoxazole proprionic acid) (non-NMDA) receptor-associated channels, which allow entry of sodium ions to depolarise the postsynaptic cell. Calcium enters the postsynaptic cell via the NMDA channel only when the cell is sufficiently depolarised to relieve the voltage-dependent block by Mg^{2+} (such as during tetanus in experimental situations) and bound to glutamate. Thus the NMDA receptor-associated channels are gated by both ligand and voltage, providing a specific window of conditions for activation.

(b) Maintenance of LTP

The maintenance or expression of LTP, that is the alterations which take place at the synapse to maintain the potentiated state, are a consequence of NMDA receptor activation and calcium entry. The potential targets of calcium in the dendritic spine 'compartment' of the postsynaptic cell are numerous, suggesting the involvement of a number of specific signalling pathways. Ca^{2+}/phospholipid-dependent protein kinase (PKC) is activated during the induction of LTP and remains active during the maintenance phase, independent of calcium and diacylglycerol (Klann *et al.*, 1993; Sacktor *et al.*, 1993). Such persistent autonomous activation of a kinase by autophosphorylation is consistent with a role in sustained changes at a synapse.

Ca^{2+}/calmodulin-dependent protein kinase (CaMKII), a major component of the postsynaptic density, also becomes autophosphorylated when activated by calmodulin-associated calcium ions. Once autophosphorylated, the enzyme no longer requires calcium for activity and can maintain the phosphorylated state after the initial activation (Miller and Kennedy, 1986). CaMKII was subsequently shown to be required for the induction of LTP by Malinow *et al.* (1989) by delivering a selective inhibitor of the kinase into the postsynaptic CA1 cells. This is supported by more recent experiments where gene inactivation of CaMKII results in a block of LTP (Silva *et al.*, 1992a) (see section 1.4.2(a)).

Several other kinases have been implicated in LTP although their activation as a result of calcium influx is not clear. The protein tyrosine

kinases (PTK) have been implicated through the use of specific inhibitors which, in blocking PTK activity, also block LTP (O'Dell *et al.*, 1991b). More direct evidence for PTK involvement comes from mice which lack the non-receptor PTK Fyn, where both LTP and spatial learning are impaired (Grant *et al.*, 1992) (see Section 1.4.2(a)). NMDA receptor activation has also been shown to result in the activation of MAP-2 kinase (Bading and Greenberg, 1991), which is normally activated as a part of the receptor PTK- Ras signalling pathway.

In addition to kinases, a number of other gene products are activated during LTP. The Ca^{2+}-dependent protease, calpain, has been proposed to be involved in a mechanism whereby structural changes to the synapse might sustain the potentiation (Lynch and Baudry, 1987). Calpain and its preferred substrate, spectrin, are localised in synaptic structures, and calpain inhibitors can block LTP (Denny *et al.*, 1990).

Inhibition of Ca^{2+}-dependent phospholipase A_2, (PLA_2), inhibits LTP *in vivo* and *in vitro* (Lynch *et al.*, 1989). PLA_2 activity produces arachidonic acid, a proposed retrograde messenger in the expression of LTP (see section 1.3.2). Arachidonic acid may also be converted to a number of metabolites whose exact role in LTP is not fully resolved (for review see Shimizu and Wolfe, 1990). Thus, while a number of signalling pathways are implicated in the induction of LTP, the precise nature and sequence of events which lead to structural modifications, changes in gene expression, or both, is not clear. Attempts to define upregulation of gene expression as a consequence of LTP have identified immediate early genes such as *zif268* (Cole *et al.*, 1989) and tPA, tissue plasminogen activator (Quian *et al.*, 1993).

1.2.2 LTP and memory

LTP in the hippocampus exhibits a number of properties which are relevant to a proposed role in memory. Firstly, LTP occurs in the hippocampus, a major site of learning, as well as other regions in the brain. It is induced rapidly and the increased strength in synaptic efficacy can persist for hours *in vitro* and for days and possibly weeks in the intact animal. Furthermore, LTP is described as associative (reviewed in Bliss and Collingridge, 1993) that is, as in many forms of

learning described, induction of LTP requires the association of two events, presynaptic stimulation and postsynaptic depolarisation .

It should be acknowledged, nevertheless, that while LTP is currently the best model for memory in mammals, experimental methods used to induce LTP are not likely to occur naturally in the animal. However, LTP has been produced in CA1 by very brief trains of stimulation delivered to simulate discharge patterns of hippocampal neurons in animals during exploratory situations (Larson *et al.*, 1986). Theta rhythm (4-7 Hz) is an intrinsic pattern of rhythmic oscillations observed in the rat hippocampus electroencephalogram (reviewed in Stewart and Fox, 1990). In rat, theta rhythm occurs primarily during exploratory movement, olfactory learning and REM sleep. Surgical manipulations which eliminate theta rhythm have produced spatial learning deficits in the rat similar to those obtained by hippocampus lesions (Winson, 1978).

(a) NMDA antagonists and spatial learning in rats.

Spatial learning in rats has been shown to be dependent on an intact hippocampus. Ibotenic acid-produced lesions in the rat hippocampus and anatomically related regions have supported findings in primates and humans, while amygdaloid damage failed to result in an obvious impairment (reviewed in Jarrard, 1993). The role of the NMDA receptor has also been confirmed in the rat. The NMDA antagonist D-APV (2-amino-5-phosphonovalerate) has been administered to rats and effectively blocks LTP *in vivo* in a dose-dependent manner (Collingridge *et al.*, 1983).

Perhaps the best pharmacological evidence to date for a role of LTP in learning and memory has come from rat studies. If LTP can be blocked experimentally *in vivo*, then behavioural or learning deficits might also be predicted if LTP is involved in learning and memory. Morris *et al.* (1986) applied such reasoning to a spatial learning test. In the spatial learning test, an escape platform is submerged in a fixed location within a tank of opaque water. The animal is released into the pool from variable positions and uses visual cues around the tank to navigate to the platform. In the experiments by Morris *et al*, APV was

administered intraventricularly to the rats for four days, then their performance in the spatial water maze assessed. The APV-treated animals showed an impairment in spatial learning compared to control groups, and escape latencies between the two groups were significantly different. Analysis of the swimming paths with the platform removed showed that, while the control groups spent most time in the quadrant which previously contained the platform, the experimental animals continued to swim around the extent of the pool. These pharmacological experiments provided the first experimental evidence to support a relationship between LTP and memory.

1.3 Second and retrograde messengers

1.3.1 Second messengers in LTP

While section 1.2.2 describes the requirement and consequences of ionotropic receptor activation in the formation of LTP, the involvement of metabotropic glutamate receptor (mGluR) activation has been described in the induction of LTP. Using a specific antagonist, Bashir *et al.* (1993) showed a requirement for metabotropic glutamate receptors in the induction of NMDA receptor-dependent LTP in the hippocampus. Metabotropic glutamate receptors belong to a large family of membrane-spanning proteins which principally initiate a second-messenger cascade in the cell. The metabotropic receptors are all structurally similar, spanning the membrane seven times, and several different types of second messenger systems may operate postsynaptically as a consequence of metabotropic receptor activation (reviewed in Hille, 1992).

The metabotropic receptors are coupled to G-proteins, through which activation of adenylate cyclase and the production of the second messenger, cyclic AMP (cAMP), results (reviewed in Hille, 1992). cAMP activates the cAMP-dependent protein kinase A (PKA), by binding to the regulatory subunit of the kinase. This releases the catalytic subunit, resulting in the phosphorylation of substrate proteins. PKA is involved

in the maintenance phase of LTP, as inhibitors of PKA can block the late phase of LTP, having no effect on normal synaptic transmission or the induction of LTP (Frey *et al.*, 1993). cAMP has also been shown to modulate directly channel conductance in olfactory neurons without phosphorylation (Nakamura and Gold, 1987), which is initiated by activation of the large '7-TM' family of odourant receptors.

Metabotropic receptor activation can also result in the activation of phospholipase C, which produces inositol triphosphate (IP_3) and diacylglycerol (DAG) from membrane phospholipids. IP_3 binding to its receptor triggers release of Ca^{2+} from intracellular stores, which can then activate Ca^{2+}-dependent kinases. DAG is confined to the membrane, where, together with phospholipids and calcium ions, it activates protein kinase C (PKC), which translocates from the cytoplasm to the membrane. PKC exists as ten different isoforms, seven of which are expressed in the brain, and some of which,such as PKCγ, are brain-specific (Schwartz, 1993). The major forms of PKC are activated by calcium, however Ca^{2+}-independent forms also exist and these are proposed to function in the maintenance phase of LTP (Saktor *et al.*, 1993). Activation of PKC results, for example, in the phosphorylation of NMDA receptor subunits on a number of different sites of the intracellular domain (Tingley *et al.*, 1993), and stimulation of PKC has been shown to modulate NMDA-receptor channel currents (Chen and Huang, 1992).

1.3.2 Retrograde messengers in LTP

While the induction of LTP requires the postsynaptic activation of ionotropic and metabotropic receptors and an elevation in intracellular calcium levels, the sustained efficacy of the synaptic connection is thought to require a presynaptic modification to sustain the observed increase in glutamate release (Bliss *et al.*, 1990) and synaptic vesicle mobilisation (Bekkers and Stevens, 1990). A membrane-permeable retrograde messenger, which facilitates synaptic communication from the postsynaptic cell to the presynaptic terminal, has been proposed to communicate the potentiated state to the presynaptic cell. At present the best candidates for a retrograde messenger are nitric oxide (NO) and

arachidonic acid, however other molecules such as carbon monoxide (Zhuo *et al.*, 1993) and platelet activating factor (Wieraszko *et al.*, 1993) have also been proposed.

Arachidonic acid, as well as 11- and 12-hydroxyeicosatetraenoic acids (metabolites of arachidonic acid), was shown by Dumuis *et al.* (1988) to be produced from striatal neurons upon glutamate or NMDA application. Subsequently, arachidonate was shown to be released upon production of LTP by tetanic stimulation of the perforant path (Lynch *et al.*, 1989). Furthermore, when the perforant path is stimulated weakly in the presence of arachidonic acid, a slow-onset increase in synaptic strength is observed in the dentate gyrus (Williams *et al.*, 1989b). This is accompanied by an increase in glutamate release. O'Dell *et al.* (1991a), however, repeated these latter experiments at the Schaffer collateral/CA1 synapse and found that while the synapse was potentiated, the effect could be blocked with the NMDA antagonist, APV. Such a finding is inconsistent with an exogenously applied retrograde messenger, which should act independently of the NMDA receptor.

Nitric oxide is produced from arginine by the Ca^{2+}-calmodulin-dependent enzyme nitric oxide synthase (NOS). NO is a diffusible gas with a very short half-life, making it a candidate for a retrograde messenger across the synaptic cleft. Schuman and Madison (1991) showed that injection of a specific inhibitor of NOS, N^G-methyl-L-arginine, into the postsynaptic cell inhibits LTP. In addition, perfusion of the hippocampal slice with haemoglobin, which will sequester NO released from the postsynaptic cell, also inhibits LTP. However, although a brain isoform of NOS exists, it is not expressed in the CA pyramidal cells of the hippocampus (Bredt *et al.*, 1991a). It should be noted, however, that an endothelial isoform of NOS is expressed in CA1 pyramidal cells (O'Dell *et al.*, 1994) and NOS inhibitors can still block LTP in mice which lack the brain isoform, suggesting that the endothelial NOS is responsible or compensatory for NO production in the hippocampus.

1.3.3 Arachidonic acid metabolism

As a result of transmembrane signalling, the potential retrograde messenger, arachidonic acid, can be produced from phospholipids from the plasma membrane by calcium-activated phospholipase A_2 (PLA_2), or from diacylglycerol after phospholipase C ($PLC\gamma$) and DAG lipase activities (reviewed in Piomelli and Greengard, 1990). Arachidonic acid can either diffuse out of the cell, reincorporate into the phospholipid bilayer, or undergo metabolism to a number of arachidonic acid-derived metabolites such as prostaglandins and leukotrienes. Enzymes mediating the metabolism of arachidonic acid include lipoxygenases, cyclooxygenases and cytochromes P-450.

Lipoxygenases are a group of enzymes involved in the arachidonic acid cascade which results in the formation of leukotrienes and other diverse metabolites. Leukotrienes participate in the local inflammatory response, acting as vasoconstrictors (reviewed in Rang and Dale, 1987). 5- and 12- lipoxygenases are expressed in the brain and formation of 12-lipoxygenase metabolites has been shown after NMDA receptor activation. Dumuis *et al.* (1988) demonstrated the production of arachidonic acid, as well as an increase in 12-hydroxyeicosatetraenoic acid (12-HETE), a product of 12-lipoxygenase activity, after the activation of NMDA receptors. Williams *et al.* (1989a) have shown that nordihydroguaiaretic acid (NDGA), an inhibitor of lipoxygenase activity, can attenuate the induction and maintenance of LTP.

Cyclooxygenase (COX-1), prostaglandin H (PGH) synthase, is the rate-limiting enzyme in prostaglandin and thromboxane synthesis from arachidonic acid (Yamagata *et al.*, 1993). An inducible form of cyclooxygenase (COX-2) was identified from differential screens to identify mRNAs induced upon seizure (Yamagata *et al.*, 1993). Prostaglandins are produced in the brain, as is the immediate precursor, arachidonic acid, as well as PLA_2 and PLC, the enzymes responsible for arachidonic acid production. The observed range of actions for prostaglandins as potent vasodilators is diverse (reviewed in Rang and Dale, 1987). COX-2 expression appears to be tightly regulated, being rapidly and transiently induced in the brain after seizure and NMDA-mediated synaptic activity (Yamagata *et al.*, 1993). However it should be noted that Williams *et al.* (1989a) found no effect

of indomethecin, a cyclooxygenase inhibitor, on LTP in CA1 of hippocampal slices.

Peppelenbosch *et al.* (1993) have demonstrated a role for cyclooxygenase and 5-lipoxygenase metabolites of arachidonic acid as second messengers in actin remodelling *in vitro*. Epidermal growth factor (EGF) -induced changes in fibroblast cell morphology are inhibited by the PLA_2 inhibitor, 4-bromophenylacyl bromide, implicating arachidonic acid metabolism in the morphological changes. Using the cyclooxygenase inhibitor, indomethacin, and the 5-lipoxygenase inhibitor, nordihydroguaiaretic acid, prostaglandins and leukotrienes were shown to regulate depolymerisation and polymerisation of actin respectively. Members of the Rho family of small GTP-binding proteins are known regulators of this growth factor-induced actin reorganisation (Ridley *et al.*, 1992) and interestingly, Han *et al.* (1991) have shown that prostaglandins and arachidonic acid can directly modulate that pathway. Because arachidonic acid production has been described in examples of induced synaptic activity (Dumuis *et al.*, 1988; Lynch *et al.*, 1989), the possible involvement of prostaglandins and leukotrienes in modulating synaptic efficacy is of great interest.

1.3.4 Cytochromes P-450

Cytochromes P-450 are a large superfamily of monoxygenases, members which are represented in bacteria, yeast, plants and mammals. While the roles of cytochromes P-450 are diverse they all carry out essentially the same function, that is, to render substrates less hydrophobic and, in the case of a foreign compound, more easily excreted. Cytochromes P-450 cleave molecular oxygen to water, inserting the remaining oxygen into the hydrocarbon bond of an aromatic or aliphatic molecule (reviewed in Poulos, 1988). This requires a coenzyme (NADPH-cytochrome P-450 reductase) for the transfer of electrons and a protoporphyrin IX haem group non-covalently bound to the cytochrome P-450.

According to Nelson *et al.* (1993), there are currently more than 200 *CYP* genes, which are grouped into 36 families on the basis of sequence homology. Of these, 12 families exist in mammals, which are further subdivided to comprise 22 mammalian subfamilies of cytochrome P-450.

Cytochrome P-450 genes encode membrane-anchored proteins of approximately 500 amino acids (Nelson and Strobel, 1988) and all share a number of common amino acid motifs which are not only relevant structurally and functionally to the protein, but are also diagnostic for cytochrome P-450 identification. Protein crystallographic studies of the bacterial CYP101, although not membrane-associated, have detailed functionally important domains of cytochromes P-450 (Poulos *et al.*, 1985).

There is absolute conservation of a cysteine residue towards the carboxy-terminus of all cytochromes P-450, which provides contact with the haem group (Poulos, 1988). The homology extends to amino acids surrounding the cysteine, where a fairly strict consensus sequence, FXXGXXXCXG (cysteine is amino acid 357 in CYP101) is diagnostic for a cytochrome P450 (Nelson *et al.*, 1993). The three-dimensional structure around the cysteine is conserved, providing a consistent environment for the protein-haem interaction. The O_2-binding pocket, though less-tightly conserved, can generally be predicted by the G/AGXXT motif (threonine is amino acid 252 in CYP101), where the threonine creates a local widening of the protein structure, creating a part of the O_2 pocket (Poulos *et al.*, 1988). There are a number of other conserved amino acids within different cytochrome P-450 proteins which may predict the type of function for a P-450. Thus, for example, xenobiotic-metabolising types of cytochromes P-450 are generally, though not exclusively, grouped in the same gene families (reviewed in Gonzalez, 1992).

(a) Biosynthetic pathways mediated by cytochromes P-450

Aside from xenobiotic activities of cytochromes P-450 (for review see Gonzalez, 1992), a number of cytochrome P-450 family members are involved in normal steroidogenic and catabolic activities. The three

major biochemical pathways with cytochrome P-450 involvement are steroid biosynthesis, the arachidonic acid cascade and bile acid biosynthesis. The enzymes involved in these pathways are the sole identified members of families CYP7, CYP17, CYP19, CYP21 and CYP27, and family CYP11 which has two members to date. Members of the CYP4 family, while not involved in steroidogenesis, are responsible for the hydroxylation of fatty acids and their derivatives.

The principal site for steroid hormone biosynthesis is the adrenal gland, however other tissues, including the brain, are steroidogenic. Cytochromes P-450 play a major part in the conversion of cholesterol to the various types of steroids (Figure 1.3) (reviewed by Miller, 1988). Cholesterol is derived either from low density lipoproteins (LDL), which transport insoluble cholesterol molecules through plasma to the cell, or by cholesterol biosynthesis.

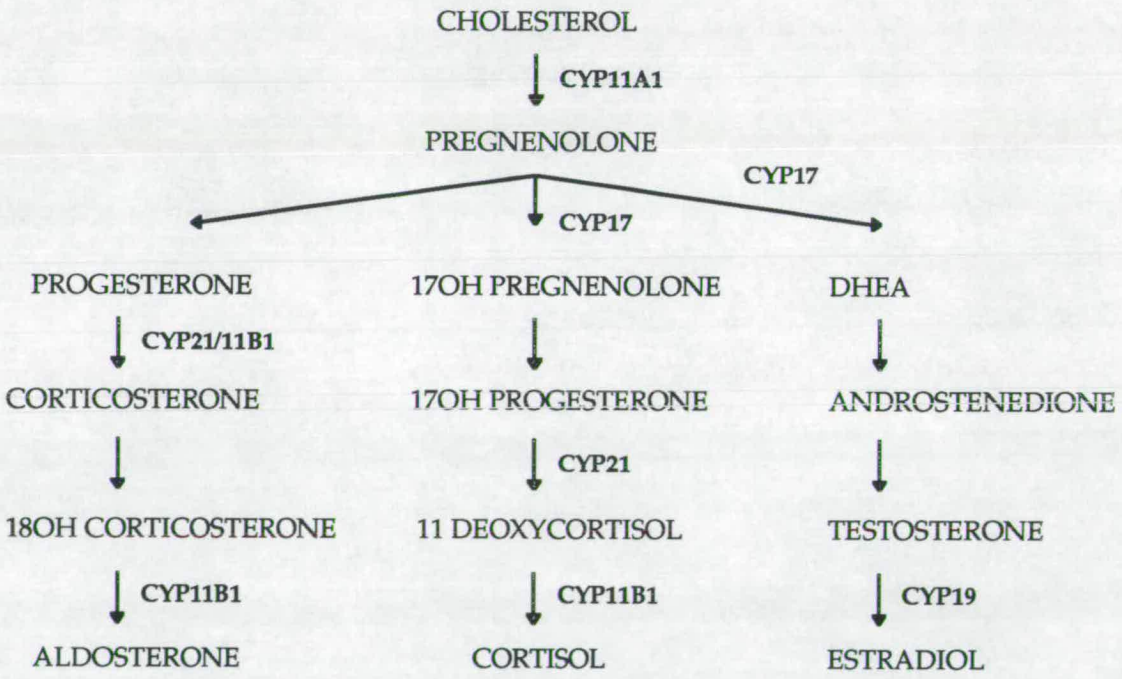


Figure 1.3 A summary of steroid biosynthesis, indicating cytochrome P-450 involvement.

The first and rate-limiting step in the synthesis of all steroid hormones is the hydroxylation of cholesterol to pregnenolone by the mitochondrial cytochrome P-450, CYP11A1, which is localised on the mitochondrial membranes (Morohashi *et al.*, 1987). Pregnenolone is then converted through one of three pathways to either glucocorticoids, mineralocorticoids or the sex hormones (Figure 1.3) through a combination of cytochromes P-450, as well as other enzymatic activities.

Arachidonic acid, released from the cell membrane upon receptor activation, is also a cytochrome P-450 substrate (Piomelli and Greengard, 1990). CYP4A1 has a rather narrow substrate specificity for the hydroxylation of arachidonic acid, as well as lauric acid and palmitic acid, and is expressed predominantly in liver (Hardwick *et al.*, 1987). A large number of metabolites are produced from arachidonic acid, and it is possible that the effects attributed to arachidonic acid in LTP (Williams *et al.*, 1989b), for example, may be due to downstream metabolites of the fatty acid. Cytochromes P-450 are also capable of modifying the lipoxygenase and cyclooxygenase metabolites of arachidonic acid. CYP4A4, for example, catalyses the ω-hydroxylation of several prostaglandins in lung (Matsubara *et al.*, 1987).

In liver, cytochromes P-450 participate in a series of reactions which result in the conversion of cholesterol into bile acids (Bjorkhem, 1985). This functions as a means of cholesterol disposal, thereby regulating cholesterol levels in the body. The principal bile acids formed from cholesterol in mammals are cholic acid and chenodeoxycholic acid, which are then converted to deoxycholic acid and lithocholic acid respectively, by microbial activity. Three cytochromes P-450 participate in the series of reactions to produce bile acids. The first and rate-limiting step is controlled by the cytochrome P-450, cholesterol 7α-hydroxylase (CYP7) (Li *et al.*, 1990). This is the first reaction of the pathway where CYP7 hydroxylates cholesterol at position seven of the cholesterol backbone. CYP27 also participates in bile acid production, by hydroxylation of the cholesterol side chain (Cali and Russell, 1991). Thirdly, the 12α-hydroxylation step, is catalysed by an enzyme for which no gene has yet been identified.

CYP7, as the rate-limiting enzyme in cholesterol degradation, is under tight regulatory control. Its transcription is subject to negative feedback activity by bile acid build-up, while cholesterol and cholestyramine, a resin that binds bile acids, exert positive control (Ramirez *et al.*, 1994). In addition, *CYP7* transcription is under circadian control in the liver, with the highest amounts of CYP7 being produced in the evening in rats (Lavery and Schibler, 1993). Such levels of regulation are not unexpected for a protein which must balance the requirement for cholesterol in membrane assembly and steroidogenesis, and the requirement for bile acids in digestion. The consequences of improper regulation of this pathway could therefore be a factor in diseases such as artherosclerosis which is associated with increased levels of circulating cholesterol (Breslow, 1993).

(b) Evidence for cytochrome P-450 activity in the brain

Specific roles for cytochromes P-450 in the brain are not fully defined. Members of both xenobiotic and steroidogenic families of cytochromes P-450 are expressed in the brain, however, suggesting that metabolism of steroids, fatty acid derivatives and exogenous compounds takes place here. Most of the studies to detect cytochrome P-450 activity in the brain have not used *in situ* methods of detection, thus the regional localisation of individual P-450s within the brain is not well characterised.

Expression of cytochrome P-450 proteins including CYP4A, CYP1 and CYP2 family members in rat brain was detected using Western analysis by Warner *et al.* (1993). In addition, these enzymes were shown to be differentially upregulated in brain under different conditions such as pregnancy and ethanol treatment. CYP19, also known as aromatase, the enzyme responsible for the conversion of testosterone to the female sex hormone estradiol, is expressed in the brain (Martini and Melcangi, 1991). Thus the effects of these sex hormones is determined by aromatase expression in the target neuron.

Distribution of the coenzyme, NADPH cytochrome P-450 reductase, as a marker of cytochrome P-450 activity, was shown by Haglund *et al.* (1984) to be restricted to catecholamine-producing neuronal

populations in rats and monkeys. However, this may not be a complete reflection of the expression pattern, as Bredt *et al.* (1991b) allude to unpublished results of the expression of NADPH cytochrome P-450 reductase in other sites within the rat brain, such as Purkinje cells of the cerebellum.

(c) Nitric oxide synthase as a combined cytochrome P-450 system.

As described in section 1.3.2(b)), nitric oxide synthase (NOS) is the enzyme responsible for the production of nitric oxide, a putative retrograde messenger in the maintenance of long-term potentiation. The brain isoform of NOS has been cloned, and amino acid sequence comparisons have shown significant homology to NADPH-cytochrome P-450 reductase (Bredt *et al.*, 1991b). This region of homology is confined to the carboxy-terminal of NOS, and includes conservation of the NADPH binding site, the cofactor necessary for electron transfer.

It has been proposed that NOS represents a complete cytochrome P-450 system (White and Marletta, 1992). NOS contains a haem group, and shows the characteristic shift in absorbance at 450 nm in reduced conditions. The presence of a haem and a reductase in the same molecule suggests that NOS may be a self-sufficient P-450 system. If so, NOS represents the first soluble P-450 in eukaryotes.

However this biochemical data is not supported by sequence comparisons. Assuming that the amino-terminal half of NOS mediates the cytochrome P-450 function, no significant homologies have been found (Nelson *et al.*, 1993). In addition, the highly conserved amino acid motifs, such as the haem-binding domain (FXXGXXXCXG) of cytochromes P-450, are not present. Crystallographic studies of NOS may shed light on this discrepancy.

1.4 Methodologies to test brain function

As described in section 1.2, strategies to understand brain function and memory have been based largely on neuropsychological studies of human amnesia and animal models. The development of long-term potentiation at the synapse as a working model for memory has led to several proposed mechanisms underlying the formation of memory in the hippocampus. The dissection of the biochemical pathways, such as those described above, and the role of specific molecules in LTP and memory can be approached through molecular intervention.

Gene replacement and gene addition technologies, where animals are created with new genetic combinations which can be transmitted through the germline, can be used to dissect the role of the hippocampus and specific molecular pathways involved in LTP and the formation of memory.

1.4.1 Addition transgenesis

Transgenesis results in the expression of extra copies of a selected gene under the control of a promoter chosen to specify spatial and temporal expression. The DNA construct is injected into one of the pronuclei of fertilised eggs, which are then reimplanted in a foster mother and allowed to develop to term. The exogenous DNA integrates randomly into the host genome, usually in multiple tandem copies, and can be transmitted to their offspring. Thus each line of animals generated is distinct in terms of transgene copy number and integration site. This technology is applicable to mice and rats, as well as several other species (reviewed in Lathe and Mullins, 1993).

In order to direct a specific pattern of expression, large pieces of DNA (up to 30-40 kb) containing the coding region and associated regulatory region(s) can be injected. As much as 650 kb, containing the entire amyloid precursor protein gene, has been introduced using a YAC (yeast artificial chromosome) vector (Lamb *et al.*, 1993). Alternatively, chimaeric genes can be created using a defined promoter and the coding region of a gene to express gene products ectopically, or restrict expression to a specific location.

Transgenesis has been used to assess the role of specific genes in the nervous system and to generate animal models of human disease.

Transgenic mice harbouring a 34 kb genomic fragment containing the complete human neurofilament heavy chain gene (NF-H) were created to assess the role of this structural protein in axons (Cote *et al.*, 1993). Neurofilaments, comprising NF-H together with light (-L) and medium (-M) subunits, are the main cytoskeletal component of axons, and therefore a component of axonal transport. Abnormal neurofilament accumulation is a prominent feature of a common form of motor neuron disease, amyotrophic lateral sclerosis (ALS). The human transgene was expressed exclusively in the nervous system, and resulted in the development of neurological abnormalities similar to those found in ALS. The neuronopathy in these mice is characterised by swellings in motor neurons and some axonal atrophy, and an impairment in axonal transport is proposed as a result of NF-H overexpression.

Transgenic technology has been used to direct expression of cellular toxins in order to selectively ablate tissues or cell types (reviewed in Bernstein and Breitman, 1989). In this approach, the function of a specific cell type or tissue can be addressed rather than that of a specific gene product. Animals expressing HSV thymidine kinase under a specific promoter can be treated with ganciclovir which, upon phosphorylation by HSV-TK, can result in selective loss of cells which express the transgene. Expression of HSV-TK without its substrate is otherwise completely harmless to the mouse. The advantage of such a system is that the investigator can choose when to ablate the cell type. Such conditional cellular ablation has been used successfully to analyse the role of follicle cells of the thyroid, and the selective removal of thyroxin hormone from circulation, by ablating non-dividing thyroid follicle cells from mice (Wallace *et al.*, 1991). Conditional cellular ablation has not been reported as yet in the nervous system but can be envisaged to be a powerful technique together with suitable promoters.

1.4.2 Homologous recombination

(a) Gene inactivation

The generation of animals with specific mutations in genes allows the analysis of the gene product's function in the intact animal. This avoids the use of pharmacological inhibitors, the exact specificity of which is not always known. Unlike classical genetics where an interesting phenotype leads to a search for the gene responsible, gene targeting asks what phenotype will result from disruption of a given gene.

Embryonic stem (ES) cells are derived from the inner cell mass of the mouse blastocyst and are totipotent, being able to give rise to all cell types (Evans and Kaufman, 1981). They can be maintained in cell culture in an undifferentiated state in the presence of the cytokine DIA/LIF (differentiation inhibitory activity/leukaemia inhibitory factor) (Smith, 1991). DNA constructs containing the mutated form of the gene (usually by inclusion of the neomycin gene as selectable marker) are electroporated into embryonic stem cells, where at low frequency homologous recombination at the genetic locus occurs. ES cell clones with the desired recombination event are injected into host blastocysts, which are then introduced into a pseudopregnant female. Resulting chimaeric mice, where the ES cells have contributed to the germline, are used for breeding to produce mice homozygous for the mutated gene (for review see Hooper, 1992).

This technique is being applied to the analysis of learning and memory in mice and a number of genes with proposed roles in LTP and memory have been disrupted. The generation of mice mutant for the neural-specific α-isoform of Ca^{2+}/calmodulin-dependent protein kinase (CaMKII) was the first gene-knockout with deficits in both LTP (Silva *et al.*, 1992a) and learning (Silva *et al.*, 1992b). No obvious structural changes to the hippocampus were observed and normal synaptic transmission was unchanged, however no increase in synaptic strength between the Schaffer collaterals and CA1 neurons was observed after tetanus in mice lacking CaMKII. Silva *et al.* went on to assay for learning deficits in the mutant mice, using the water maze spatial learning test developed for rats (described in section 1.3(c)). While the mutant mice show no lack of motivation or deficits in motor skills, the escape latency of the mutant mice is always greater than wild type littermates. In addition, if the platform is removed, wild

type mice preferentially search the quadrant where the platform was located, while mutant mice fail to do so. Thus CaMKII-deficient mice show deficits in both spatial learning and LTP.

Mice lacking a functional PKCγ gene have also been generated, with effects on both LTP and spatial learning (Abeliovich *et al.*, 1993a and b). While LTP is greatly diminished in these mice, LTP can be produced when the tetanus is preceded by low frequency stimulation. This is in contrast to the conclusions drawn from the inhibition studies described (Malinow et al., 1989). The exact role for PKCγ in LTP is thus inconclusive. The presence of seven other forms of PKC expressed in neuronal tissue (Schwartz, 1993) may compensate for the loss of PKCγ, perhaps after activation by low frequency stimulation. In addition, Saktor *et al.* (1993) propose that the generation of a constitutively active PKC, the ζ isoform, after translocation to the membrane, is required for LTP maintenance.

Because protein tyrosine kinase inhibitors are known to block LTP (O'Dell *et al.*, 1991), Grant *et al.* (1992) analysed mice with null mutations in the non-receptor protein tyrosine kinases *fyn*, *yes*, *abl* and *src*. Each kinase is expressed in the hippocampus, which together with the inhibition studies, suggested a possible role in LTP for one or more of the tyrosine kinases. LTP in CA1 of hippocampal slices was the same as wild type for *yes*, *src* and *abl* mutant mice, however there was a reduction of LTP in the *fyn* mutant mice. In addition, the escape latency for the mutant mice was increased in the spatial water maze test, again confirming a link between LTP and memory. A direct role for Fyn in spatial learning and LTP is not conclusive, however, because the *fyn* mutant mice have alterations in the cellular organisation of the hippocampus. The arrangement of the granule cells in the dentate gyrus and the pyramidal cells of CA3 appears as an undulation in the structure possibly as a result of increased cell numbers in these structures. This developmental defect complicates the interpretation of the phenotype, as developmental effects cannot be separated from adult phenotypes.

Mice which lack the cytokine LIF (leukaemia inhibitory factor) also appear to have selective damage to the hippocampus. LIF is a pleiotropic factor with a wide range of activities including maintenance

of undifferentiated embryonic stem cells and cholinergic neuron differentiation (reviewed in Hilton, 1992). Initial phenotypic analysis of the mutant mice revealed no overt abnormalities apart from a failure of blastocyst implantation into the uterine wall of female mice lacking LIF (Stewart *et al.*, 1992). A preliminary abstract report claims that mice lacking LIF have several changes to the structure of the hippocampus (Patterson, 1994). There is extensive loss of pyramidal neurons, as well as decreased staining for parvalbumin, suggesting a loss of interneurons in the hippocampus. This phenotype is similar to what might have been predicted using a tissue-specific promoter directing expression of a cellular toxin (described in 1.4.1). The developmental effects of this phenotype must be considered, however, these mice potentially represent an interesting sub-region lesion of the hippocampus, worthy of electrophysiological and behavioural analysis.

(b) Gene replacement

Homologous recombination provides the means not only to delete a gene from a mouse, but also to replace the endogenous gene with a mutant version of the gene. This approach can provide more information about the gene product itself and may also provide a more subtle phenotype, especially in cases where gene inactivation results in embryonic lethality.

Hasty *et al.* (1991) developed a strategy based on homologous recombination called "hit and run" to create such subtle mutations in the mouse genome. This technique has been used to create a constitutively active tyrosine kinase, Hck, in ES cells (Ernst *et al.*, 1994). Phosphorylation negatively regulates the activity of Hck, and thus by mutating a conserved tyrosine residue in the cytoplasmic domain to phenylalanine, the role of Hck in the maintenance of the undifferentiated state of ES cells could be addressed. ES cells expressing constitutively active Hck were less dependent on LIF to prevent differentiation, suggesting that signalling through the IF receptor complex is at least in part mediated through Hck. The creation of transgenic mice carrying this mutation will facilitate further analysis of

Hck in the haemopoietic system, the major site of *hck* expression (Holtzman *et al.*, 1987).

(c) Tissue-specific gene targeting

As alluded to earlier, the consequences of gene inactivation on other parts of the animal requires consideration when behavioural and learning deficits are being analysed. Non-specific effects like motivation, for example, may be affected in the mutant mice, and this will affect performance in a particular task. Also memory and learning are not restricted to the hippocampus *per se*. Connections between the neocortex and the hippocampus are required for declarative learning (Section 1.1.2(b), (c)) and the consequences of gene activation in these regions warrants consideration. Finally, inactivation of genes such as the NMDAR1 receptor subunit which results in perinatal lethality (Li *et al.*, 1994), would more likely be of considerable value if the gene were inactivated in the adult hippocampus alone.

The use of site-specific homologous recombination systems, combined with tissue-specific promoters, has recently been demonstrated to produce gene inactivation in a chosen site of gene expression. Gu *et al.* (1994) have applied site-specific recombination, using the Cre-*lox*P system (Sauer and Henderson, 1988), to generate mice which lack the DNA polymerase β gene in T cells only. Cre is a site-specific recombinase from the bacteriophage P1, which will bind *lox*P sequences, and delete the intervening DNA, leaving one *lox*P site. The *lox*P sequence is composed of two 13 bp inverted repeats separated by an 8 bp spacer region. Cell type-specific gene inactivation is achieved by the generation of two transgenic mice: the first contains *lox*P sites flanking the DNA to be deleted in all cell types, and the second is a conventional transgenic mouse, with a *cre* transgene under the control of a tissue-specific promoter. Interbreeding will result in mice where recombination between the *lox*P sites will occur only where Cre is expressed.

1.4.3 Limitations of transgenesis

There are a number of limitations associated with both knockout technology and addition transgenesis. Addition transgenesis relies on the existence and characterisation of promoters which are capable of producing the desired spatial and temporal pattern of transgene expression and at sufficient levels. While promoters do exist to direct expression to some specific cell types, as yet no promoters have been characterised which will direct transgene expression specifically to the hippocampus.

Knockout technology also lacks specificity in that there is no selective loss of gene expression in the animal: every cell in the homozygote is affected by the deleted gene, and at all stages of development. Thus phenotypic interpretation can be complicated by the perhaps subtle deficits incurred during development. Cell type-specific gene inactivation circumvents some of the limitations associated with conventional knockout technology (see section 1.4.2(a)). However, the application of such technology to the hippocampus is entirely dependent on the identification of promoters capable of directing specific transgene expression.

The interpretation of the phenotypes generated by gene-targeting do not prove straight forward in all cases. Genetic redundancy, where the function of the targeted gene is compensated in part or in full by a related gene, has been used to explain the lack of overt phenotype in a number of knockouts. Increasingly, it is becoming clear, mainly through PCR-based approaches, that many genes are members of gene families. The lack of overt or predicted phenotype in a number of knockout mice suggests the existence of a related gene which can compensate for its loss. The fibroblast growth factor family, for example, consists of nine related proteins, which may compensate for the lack of phenotype in the FGF-5 knockout mouse (apart from sustained growth of fur) (reviewed in Mason, 1994).

Conversely, the analysis of genes which are essential for viability is limited. The NMDAR1 subunit, in combination with other NMDA receptor subunits, is required for a functional NMDA channel (Monyer *et al.*, 1992), and activation of the NMDA receptor-associated channel is the first event in the induction of LTP. Thus the role of the individual NMDA receptor subunits in the context of hippocampal LTP is of

interest, but cannot be addressed in the knockout mice, as homozygotes die shortly after birth (Li *et al.*, 1994).

Finally, the remaining limitation to knockout technology is that at present no ES cells have been derived from species other than mouse. Thus knockouts in rats, for which much of the learning and memory work was done previously, cannot complement such studies.

1.5 Gene expression in the hippocampus

An investigation of gene expression in the hippocampus would be a complementary avenue of investigation to the above-described research. Our understanding of the role of the hippocampus in memory and learning can be expanded by the application of transgenesis and gene-targeting. Indeed, as described (section 1.4.2), gene-targeting has already made substantial progress in elucidating the roles of kinases and phosphorylation in memory and LTP. However, modifications directed specifically to the hippocampus would add a level of refinement to such experimental approaches. In addition, hippocampus-specific gene inactivation requires expression of Cre directed specifically to the hippocampus.

An investigation of gene expression in the hippocampus is a prerequisite for such experiments. A knowledge of which genes are expressed in the hippocampus, and the exact spatial and temporal distribution of their products will provide markers to distinguish different cell types in the hippocampus. The function of such genes in terms of synaptic plasticity and long-term potentiation can also be addressed. Finally, if the expression of such genes is restricted to the hippocampus or subregions thereof, then their transcriptional regulatory elements can be exploited to direct gene expression to the hippocampus. With such promoters conditional ablation of the hippocampus or subregions could define the role of specific regions of the hippocampus in memory. In addition, over-expression of candidate regulatory molecules, dominant-negative variants or other suitably modified versions specifically in the hippocampus can be used to assess

the specific role of a gene product in the biochemistry of LTP and memory.

CHAPTER TWO

IDENTIFICATION OF MOLECULAR CLONES ENCODING HIPPOCAMPUS-SPECIFIC TRANSCRIPTS

Differential hybridisation has been widely used to identify cDNA clones encoding transcripts whose expression is restricted to particular tissues or stages of development (Sambrook *et al.*, 1989, and references therein).

cDNA clones expressed in a particular cell or tissue type can be isolated using such techniques as differential hybridisation. cDNAs of interest are isolated by screening a cDNA library with a probe made from mRNA from the tissue of interest. At the same time, the library is screened with a probe made from mRNA from a tissue that does not contain the sequences of interest. Clones which hybridise with the first probe and not the second are purified for further analysis.

More refined methods include enrichment steps for either the library or the probe to remove selectively common or highly repeated sequences. This has been used in the case of subtracting libraries where the floor plate-specific F-spondin was isolated (Klar *et al.*, 1992). The myogenic determination factor, MyoD, was isolated using subtracted cDNA probes (Davis *et al.*, 1987). Both differential and subtractive hybridisation increase the probability of isolating clones of interest.

To identify transcripts whose expression is enriched in or restricted to the hippocampus, a differential hybridisation screen was performed. The rat was chosen because most behavioural analysis and hippocampus lesions have been performed with the rat. In addition, the rat hippocampus allowed for greater ease of dissection and provided more starting material. Rats were sacrificed and the hippocampus removed with no obvious contamination from other brain regions; the remainder of brain ("ROB") was similarly excised. For the preparation of the hippocampus library, these dissections were

performed by Prof. R. Morris, Dept. Pharmacology. PolyA$^+$-selected RNA was purified and used to prepare cDNA libraries in *lambda* ZAP II by Dr. J. Mason.

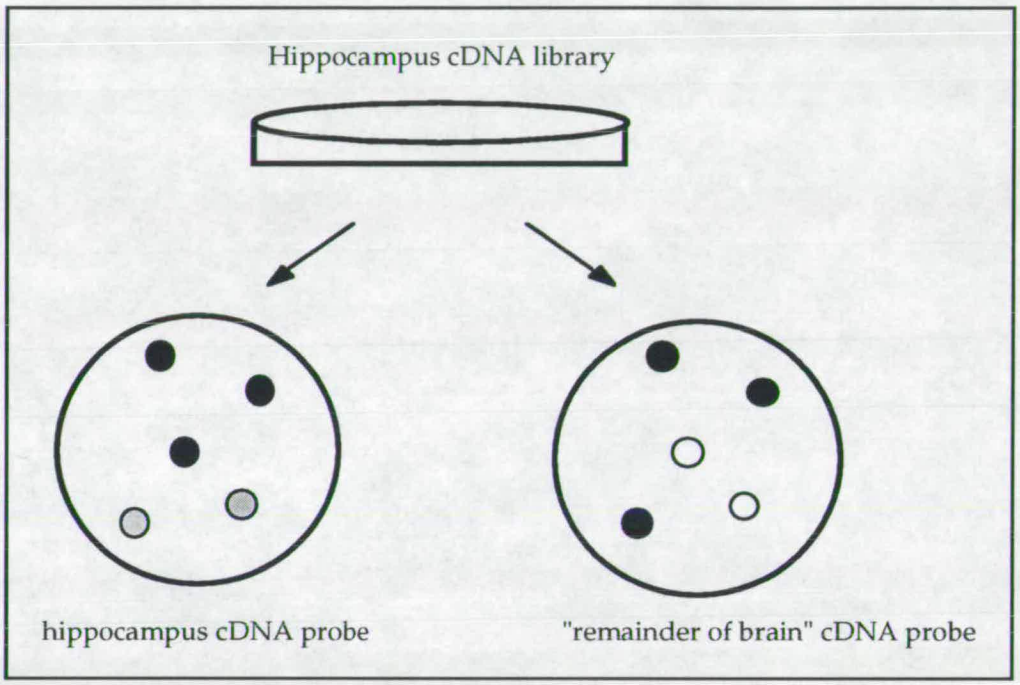


Figure 2.1: Differential hybridisation of a hippocampus cDNA library.

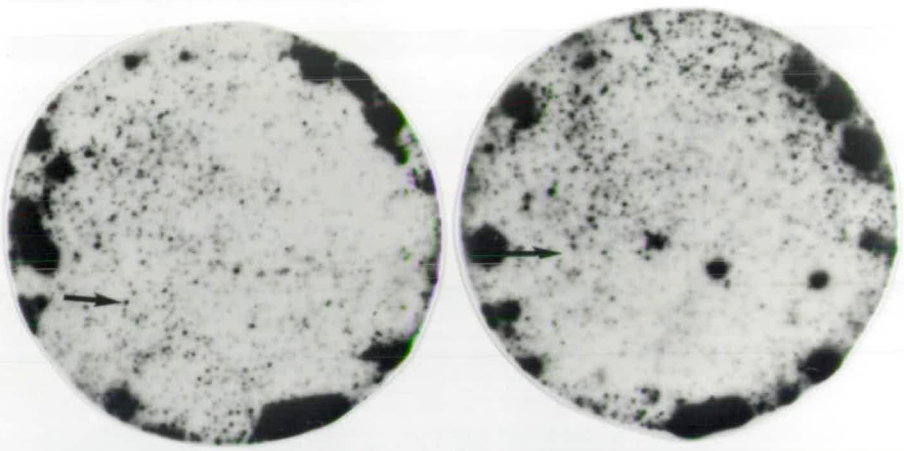
2.1 Differential screening of a rat hippocampus cDNA library

A total of 10^6 recombinant clones from the hippocampus cDNA library was plated out at a density of 50,000 recombinant phage per 15 cm plate. Duplicate lifts of the library were hybridised to cDNA probes prepared from reverse transcribed RNA from the hippocampus and "ROB". After overnight hybridisation, the filters were washed and exposed for autoradiography (for example, see Figure 2.2).

From this primary screen, 361 recombinant phage plaques gave a differentially hybridising signal and were picked. The intensity of the signals varied, possibly indicating relative abundance of the transcripts.

Figure 2.2 Differential hybridisation filters. (a) a representative filter from the primary screen, and (b) the secondary screen. For both panels, filters on the left hand side were probed with radiolabelled hippocampus cDNA, and the filters on the right side were probed with a cDNA probe prepared from "remainder of brain" after hippocampus removal. Exposure times in these examples were 2 days for the primary filters and 17 hours for the secondary filters.

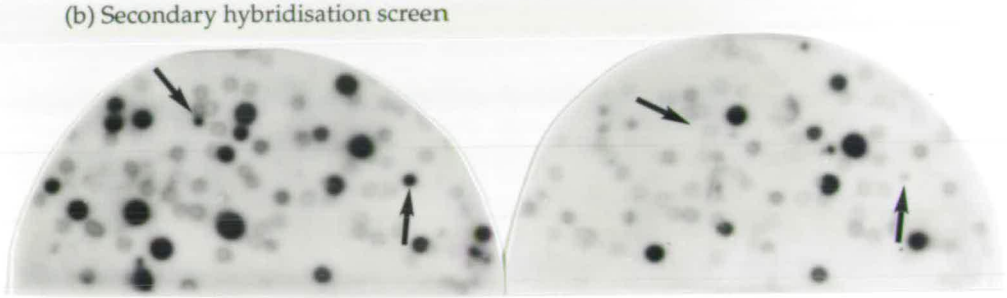
(a) Primary hybridisation screen



hippocampus cDNA probe

"remainder of brain" cDNA probe

(b) Secondary hybridisation screen



hippocampus cDNA probe

"remainder of brain" cDNA probe

The differential signals were dependent on exposure time, and this was optimised so that the greatest differences between duplicates could be visualised. Of the 361 clones, 75% of these gave no signal on the negative filter, while the remaining 25% gave a very faint signal with the "ROB" probe. This latter group might represent highly enriched transcripts in the hippocampus with perhaps restricted expression elsewhere in the brain. Of the former 75%, again the signal intensity varied: 10% of the clones gave a very strong hybridising signal, 51% gave a 'medium' signal intensity, and 39% gave a very weak signal. Again this may reflect the relative abundance of different transcripts within the hippocampus.

As an initial step, a representative sample of 49 differentially hybridising plaques was characterised further. Each recombinant was purified by secondary screening using the same probes used in the initial screen. The cDNAs were then rescued from each *lambda* ZAP II phage as pBluescript II SK plasmids using the ExAssist/SOLR system (Stratagene), and partially sequenced from each end using primers hybridising to the T3 and T7 polymerase binding regions which flank the cloning site. Sequence data were compared with the GenBank/EMBL sequence database, using the Fasta programme (UWGCG version 7.0) to detect previously described transcripts. The outcome of this search is summarised in Table 2.1.

Of the 43 cDNA clones selected, 9 were identified in the database (21%), 11 contained repeat sequences (25%), 2 contained plasmid sequences, and the remaining 21 were novel (49%).

The average size of the inserts in the hippocampus cDNA library is approximately 1.0 kb (personal communication, Dr. J. Mason). Because the library was oligo-d(T)-primed, the majority of sequence analysed would be derived from the 3'UTR of each transcript, complicating clone identification. Rat orthologues of already identified genes from other species may not be identified unless the 3' untranslated region was highly conserved between species. Similarly, structural and functional motifs contained within the coding sequence, such as DNA binding or trans-membrane domains would not be represented in partial length clones.

clone number	insert size (kb)	primary hybridisation signal	sequence identity	reference
3.28a	0.3	+++++	mouse PDGFR (91%, 127 bp)	Yarden *et al.*, 1986
4.5a	0.35	+++	mitochondrial gene	Gadaleta *et al.*, 1989
6.19d	5.0	++++	β-adaptin	Kirchhausen *et al.*, 1990
9.9a	1.2	+++++	calreticulin	Murthy *et al.*, 1990
11.2a	2.1	++	mouse ECA39 (72%, 161 bp)	Niwa *et al.*, 1990
13.9c	0.6	++	mouse ribosomal protein S13 (93%, 231 bp)	MacMurray and Shin, 1990
14.4a	2.0	++++	calmodulin RCM3	Nojima *et al.*, 1987
14.15a	1.1 + 1.5	++	APP	Shivers *et al.*, 1988
16.26b	1.3	++++	mouse SNAP (98%, 123 bp)	Oyler *et al.*, 1989

Table 2.1 Sequence analysis of hippocampus cDNA clones.

Of the identified cDNAs, all were reported as high abundance mRNAs. Rat amyloidogenic glycoprotein (rAG) is the orthologue of the human amyloid precursor protein (APP) which is associated with the neuropathology of Alzheimer's disease. The adult expression pattern of rAG mRNA in the brain is widespread and abundant, with the hippocampus one of the highest regions of expression (Shivers *et al.*, 1988). SNAP-25 (synaptosomal-associated protein) is a protein of as yet unknown function, but is localised in the presynaptic regions of specific but diverse populations of CNS neurons. Transcripts for SNAP-25 are enriched in the mossy fibres and the molecular layer of the dentate gyrus as well as several other areas of the brain (Oyler *et al.*, (1989) (neocortex, piriform cortex, anterior thalamic nuclei, pontine nuclei, and granule cells of the cerebellum.) The identification of these three hippocampus-enriched transcripts indicated that the differential screening procedure had been at least partly successful.

However, the other identified cDNA clones showed no enrichment for the hippocampus. β-adaptin is a component of clathrin-coated vesicles that mediates constitutive transport through the Golgi apparatus (Kirchhausen *et al.*, 1989). In the mouse, the receptors for platelet-derived growth factor (PDGF-R) are expressed throughout the brain as are the ligands PDGF-A and PDGF-B (Sasahara *et al.*, 1991). Calmodulin is a calcium binding protein, which together with calcium are involved in the regulation of a variety of enzyme systems. The RCM3 transcript is derived from the CaMII gene, and expression as measured by Northern analysis is highly enriched in the brain (Nojima *et al.*, 1987) Calreticulin is a calcium binding protein of the endoplasmic reticulum lumen (Murthy *et al.*, 1990). ECA39 is a transcript found overexpressed in a teratocarcinoma cell line which, upon differentiation, down regulates (Niwa *et al.*, 1990). In the mouse, ECA39 expression has been reported in brain and kidney, however no function is yet known for the gene product.

Clone 12.10a, although not listed in Table 2.1, showed no significant homology to previously identified sequences in the database, however further sequencing of longer 12.10a cDNA clones revealed a high

degree of homology with human fetal liver tropomodulin (Sung *et al.*, 1992) (Figure 2.3).

```
clone 12.10a      1 KVNKTLKSLNVESNFITGAGILALVEALRENDTLTEIKIDNQR. . . .  43
                    I  II. II. IIIIIIIII. IIIII  IIIII.  I. . I. I: IIIII.
tropomodulin    249 KENKVLKTLNVESNFISGAGILRLVEALPYNTSLVEMKIDNQSQPLG 295
```

Figure 2.3 Alignment of translated partial sequence of clone 12.10a with human tropomodulin. Partial sequence from the 5' end of 12.10a clones revealed nucleic acid homology with human tropomodulin. The clone 12.10a open reading frame was used for an amino acid alignment using the GAP programme (UWGCG).

Tropomodulin is a regulatory protein associated with the membrane skeletal network, and functions to regulate tropomyosin-actin interactions by binding to tropomyosin and thus inhibiting association with actin. Given the high degree of amino acid identity between clone 12.10a and human tropomodulin in the short stretch of identified sequence (Figure 2.3), it is possible that clone 12.10a represents the rat orthologue of tropomodulin.

Eleven clones showed homology to repeat elements in the database. Repeat elements are short interspersed sequences throughout the genome, which are estimated to comprise 20-30% of the genome (Schmid and Jelinek, 1982). The function of these elements in the genome is not known, however repeat elements are associated with RNA polymerase II transcription. Thus, Northern analysis of these clones resulted in a black smear (see clone 11.2a, Figure 2.4). The cDNAs containing these repeats are still of potential interest, however no attempt to remove the repeat region and further characterise these clones was made.

Each of the twenty-six novel hippocampus cDNA clones was screened preliminarily by Northern analysis to detect a corresponding

transcript and those with enriched expression in the hippocampus. Radiolabelled probes prepared from the novel cDNAs were used on Northern blots with 1 µg of hippocampus mRNA and 10 µg of "ROB" total RNA.

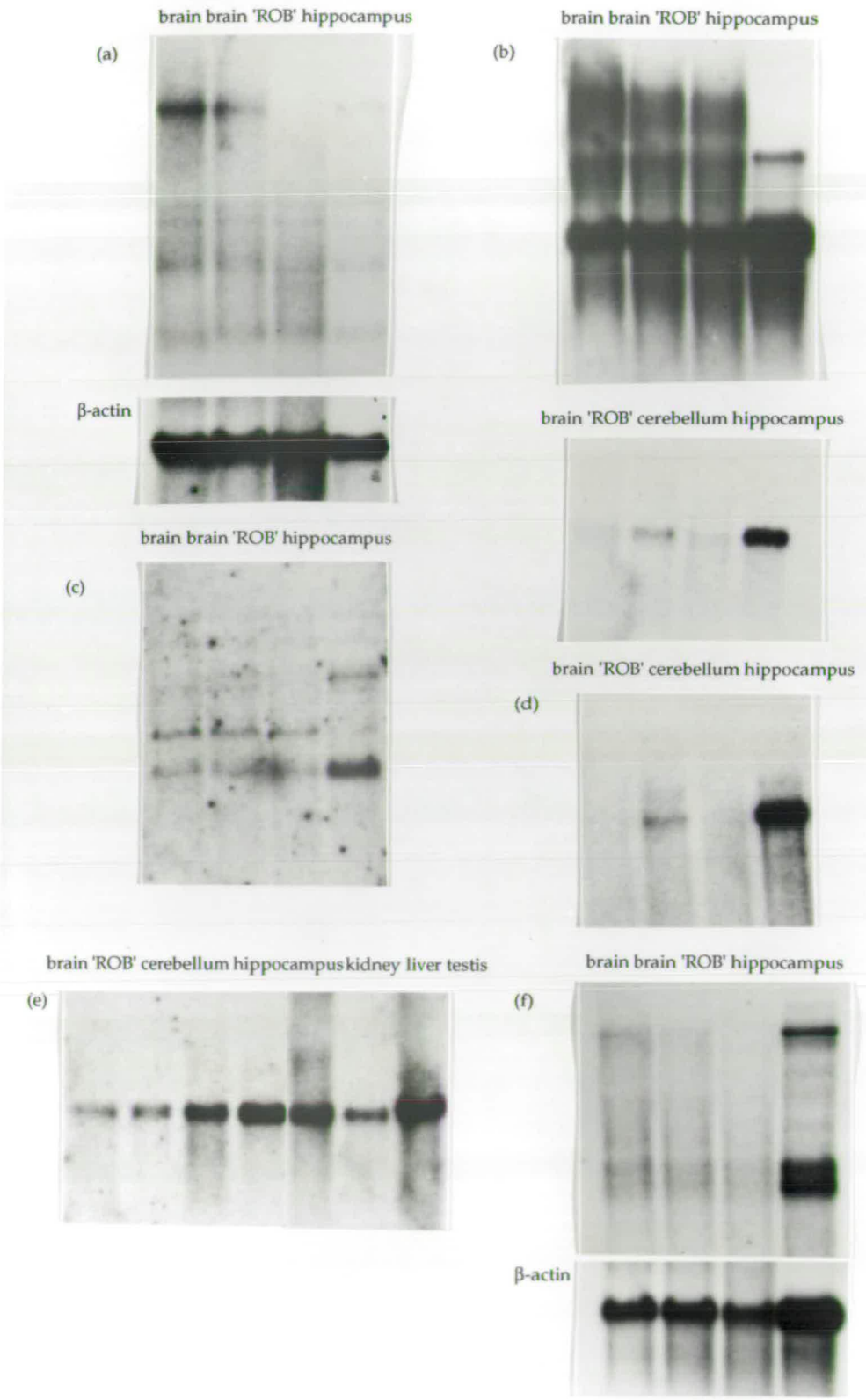
CLONE NUMBER	TRANSCRIPT(S) (kb)	HIPPOCAMPUS EXPRESSION
3.22a	4.8	enriched
3.3a	2.4	ubiquitous
4.5b	1.6	ubiquitous
11.2a	2.9	enriched
12.10a	7.0; 3.9; 3.5; 2.7; 1.9	7.0 kb 'specific'
14.4b	1.6	ubiquitous
14.5a	5.0; 2.2; 1.8	enriched
14.15b	2.25	'specific'
15.13a	4.4; 3.2; 2.45	4.4 kb 'specific'
16.26c	1.8	enriched

Table 2.2 Northern Analysis of unidentified hippocampus cDNAs.

Northern analysis identified six clones which showed transcripts enriched in the hippocampus relative to "ROB". Several cDNA clones gave no signal on a Northern blot. These clones might correspond to very rare transcripts, where Northern analysis is not sensitive enough to permit their detection. It is noteworthy that a recently described cDNA from rat brain, MR22 (also known as 5-HT$_{5B}$ (Wisden *et al.*, 1993)), which encodes a novel serotonin receptor, was only detected by

Figure 2.4 Primary Northern analysis of positive clones from the hippocampus differential hybridisation screen. The cDNA clone inserts were excised from the plasmid vector, radiolabelled by random primer, and used as probes on RNA blots. 1 μg of polyA-selected hippocampus RNA was used; all other samples were 10 μg of total cytoplasmic RNA.

(a) clone 12.10a (3 day exposure) (lower blot reprobed with β-actin), (b) clone 11.2a (upper blot was probed with entire 2.1 kb clone (1 day exposure); the lower blot was probed with a 0.1 kb 3' piece of the clone (4 day exposure)), (c) clone 15.13a (7 day exposure), (d) clone 3.22a (7 day exposure), (e) clone 14.15b (7 day exposure), (f) clone 14.5a (7 day exposure) (lower blot reprobed with actin). Intensifying screens were used for the autoradiographic exposure times indicated.


brain brain 'ROB' hippocampus
(a)
β-actin
brain brain 'ROB' hippocampus
(b)
brain 'ROB' cerebellum hippocampus
brain brain 'ROB' hippocampus
(c)
brain 'ROB' cerebellum hippocampus
(d)
brain 'ROB' cerebellum hippocampus kidney liver testis
(e)
brain brain 'ROB' hippocampus
(f)
β-actin

Northern using 5 μg of polyA$^+$ RNA and autoradiography for 2-3 weeks (Erlander *et al.*, 1993). These sequences could also be derived from non-coding sequences which contaminated the original RNA used for the library preparation.

Figure 2.4 illustrates the hippocampus enrichment that was observed at this preliminary stage of analysis. This analysis was performed using a hippocampus polyA$^+$-selected RNA sample, while the other samples were prepared as total cytoplasmic RNA. This proved to be misleading as the relative loading of samples was always biased in favour of the hippocampus sample. The ultimate assay, however, for hippocampus specific expression was to be *in situ* hybridisation.

Of the clones analysed, clone 14.5a was chosen for further characterisation, as concurrent analysis of this clone (see Chapter 2) indicated that it offered the most restricted expression pattern.

Subsequent analysis of the expression pattern of the remaining clones in tissues other than brain revealed less restricted patterns of expression. Clone 14.15b is expressed in testis, liver and kidney by Northern analysis (Figure 2.4(e)). Ongoing *in situ* hybridisation analysis revealed expression of clone 11.2a throughout the brain at high expression levels (*in situ* hybridisation data not shown), as well as non-restricted expression patterns of clones 12.10a, 15.13a and 3.22a in mouse (personal communication, Dr. M. Steel).

2.3 Attempts to characterise the remaining primary clones

From the "mini-analysis" of the differential screen, indications were that cDNAs had been isolated with some enrichment seen in the hippocampus. A large scale screening method of the remaining primary isolates was used in an attempt to eliminate ubiquitously expressed cDNAs.

Two different strategies were used to screen for hippocampus selectivity. Firstly, each of the primary isolates was dot-blotted onto a membrane in duplicate. Duplicate filters were then hybridised with either a hippocampus cDNA probe or the "ROB" probe. The dot-blot method was unsuccessful as no differential signal was observed after

several attempts. Each primary isolate is a mixture of up to 10 clones, with only one contributing to a differential signal. Thus any differential signal using the dot-blot method was most likely being concealed by the hybridisation from the other clones.

To circumvent this problem, a PCR-based method described by Mutchler *et al.* (1992) was used where the T3 and T7 primer sites flanking the insert are used to PCR amplify each insert in the mixture of clones. These amplified inserts are then fractionated on an agarose gel in duplicate, Southern blotted, and probed with either the hippocampus cDNA probe or the "ROB" probe. This approach was also unsuccessful. Each of the bands hybridised with the same intensity, regardless of the probe (data not shown). This may indicate that there were in fact no differentially hybridising cDNA clones in the primary isolates. This seems unlikely, given the positive outcome of the prescreening. Alternatively, the complexity of both the primary isolates and the cDNA probes together may not allow detection of differential signals. In addition, inserts amplified may comigrate, potentially masking differential hybridisation signals.

2.4 Discussion

A differential hybridisation screen was used to isolate transcripts whose expression is restricted to the hippocampus. Of 361 primary mixed isolates, 48 were analysed further, leaving six whose expression patterns showed apparent hippocampus enrichment.

The transcript complexity in the brain is higher than in any other tissue (Sutcliffe, 1988). Of the estimated 100,000 genes in the mammalian genome, it is estimated that approximately one third are found in the brain, and of these, 20,000 are brain specific. Thus the probes used for the screen were highly complex, and low abundance messages would be biased against in this screen. This is reflected in the genes identified in the screen, which are all expressed at high levels.

Constitutively expressed genes such as ribosomal, mitochondrial and general housekeeping genes should not be represented after a differential screen. However several such genes inluding β-adaptin, calreticulin, and mitochondrial and ribosomal genes were detected. It is

unclear why they should give a differential signal, but in other differential screens described in the literature, a small percentage of non-specific transcripts was isolated in most cases. One possible explanation is that, while constitutive, levels of these transcripts vary in different tissues or under certain conditions (Adams *et al.*, 1992).

Differential hybridisation utilizes no enrichment steps. Because of the transcriptional complexity of brain, enrichment procedures could be included in either library preparation or probe preparation. Thus hippocampal libraries or probes which have been subtracted against the remainder of brain, possibly through several rounds, will require less screening and may yield higher hippocampal specificity.

Differential display is a relatively new technique which may allow access to very low abundance hippocampus transcripts (Liang and Pardee, 1993). Differential display uses PCR to amplify sequences, using a primer pair where one is anchored to the polyA tail, the other being short and random in sequence. The differentially amplified products from different mRNA sources may be visualised by electrophoresis and subsequently cloned.

While the use of cDNA screening was employed for the isolation of spatially restricted genes, these strategies may also be applied to different induced or treated states of the hippocampus. This approach has previously been used for the isolation of seizure-induced gene expression (Qian *et al.*, 1993), and for 'candidate plasticity genes' in the hippocampus after kainate injection (Nedivi *et al.*, 1993).

CHAPTER THREE

HIPPOCAMPUS-ENRICHED EXPRESSION OF cDNA CLONE Hct1

3.1 Introduction

As a candidate hippocampus-specific cDNA, clone Hct1 (14.5a) was chosen for further expression analysis. Northern blot analysis was used to examine the adult expression pattern in other tissues of the rat, while *in situ* hybridisation was used to reveal the regional distribution of Hct1 mRNA in the adult rat brain.

The expression pattern of Hct1 in the adult mouse brain was also investigated. As embryonic stem cells are as yet only available from the mouse, any transgenic experiments utilising homologous recombination and gene targeting, would employ the mouse Hct1 genomic DNA. It was therefore necessary to confirm the expression pattern of Hct1 in the mouse brain.

3.2 Expression of Hct1 mRNA in the adult rat.

The rat hippocampus cDNA library used in the original differential hybridisation screen was rescreened to isolate longer cDNA clones of Hct1. As with the differential screen, 10^6 recombinant phage clones were plated at a density of 50,000 clones per 15 cm dish, and duplicate lifts of each plate were hybridised with a random-primer labelled probe of the 300 bp cDNA. (The analysis of Hct1 cDNA clones will be described in Chapter Four.) The longest clone isolated was 1.4 kb in length, clone 14.5a(12), and this was used as a probe in Northern blot experiments.

Total cytoplasmic RNA from various rat tissues was prepared by homogenisation in guanidinium isothiocyanate followed by purification by ultracentrifugation through caesium chloride. RNA from hippocampus was either polyA$^+$ selected RNA or total

cytoplasmic RNA. RNA was also prepared from discrete regions of the brain, such as cerebellum and olfactory bulb, which were easily dissected away from other contaminating tissue. The RNA was fractionated by denaturing gel electrophoresis and transferred to a membrane by Northern blotting.

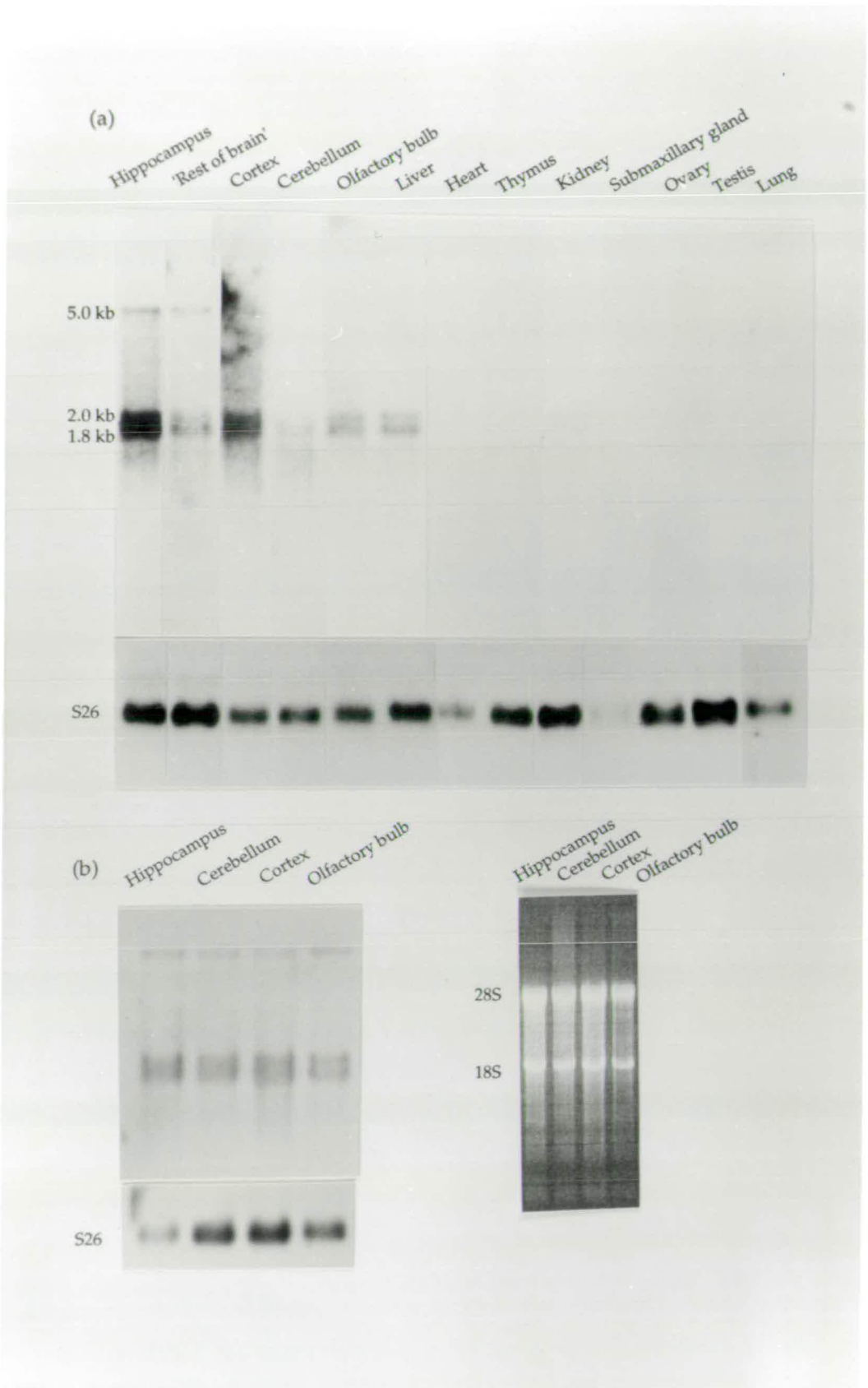
Equivalent sample loading for initial Northern blot experiments proved difficult as quantitation by optical density was not always consistent with the intensity visualised by ethidium bromide staining. In addition, initial Northern analysis was performed using polyA$^+$-selected mRNA for the hippocampus sample only, while all other samples were total cytoplasmic RNA. Equal loading of samples in this situation was thus not possible. To circumvent this problem, standardisation of RNA loading was determined by relative hybridisation to a cDNA probe for ribosomal protein S26. S26 has been described as a reliable internal control, as RNA levels are reported to be invariant in a range of eukaryotic cells and tissues (Vincent, 1993).

Probe 14.5a detected three transcripts of 1.8 kb, 2.2 kb, and 5.0 kb by Northern blot (Figure 3.1(a)). These transcripts were largely restricted to the brain. Low expression was detected in the liver, and even lower levels in the kidney, but this was always the two smaller transcripts. The 5.0 kb transcript appeared to be present only in the brain samples. This may be because the level of this mRNA in other tissues is below the level detectable in Northern blots. All other tissue types tested were negative for Hct1 expression by Northern analysis. Within the brain, the hippocampus had the highest level of all three Hct1 transcripts, however lower levels were found in all other regions tested, that is, the cerebellum, the cortex and the olfactory bulb. On re-examination of the ethidium bromide-stained gel, the cortex sample was grossly overloaded relative to other samples, despite the relative S26 hybridisation. Subsequent RNA preparations of cortex total RNA were made which did not give this result. In Figure 3.1(b) all samples were prepared as total cytoplasmic RNA and loaded so that ethidium bromide staining of the 18S and 28S ribosomal subunits was equivalent. The cortex sample now showed no significant enrichment. In addition, Hct1 signal in the hippocampus was not enriched, however the relative hybridisation of S26 was also lower, suggesting differential expression levels of S26 ribosomal protein in the hippocampus.

Figure 3.1 Northern analysis of Hct1 expression in various tissues of the adult rat.

(a) Expression levels of Hct1 message, when RNA loading was standardised using hybridisation to the internal control ribosomal protein S26 mRNA. Autoradiographic exposure was for seven days.

(b) Relative Hct1 expression levels in sub-regions of brain (RNA loading was standardised by ethidium bromide staining of the 18S and 28S ribosomal RNA subunits). Autoradiographic exposure was for four days. Approximately 10 μg of total cytoplasmic RNA was used for all tissues except the hippocampus sample in (a) where 1 μg of polyA$^+$-purified RNA was used.


(a)
Hippocampus
'Rest of brain'
Cortex
Cerebellum
Olfactory bulb
Liver
Heart
Thymus
Kidney
Submaxillary gland
Ovary
Testis
Lung
5.0 kb
2.0 kb
1.8 kb
S26
(b)
Hippocampus
Cerebellum
Cortex
Olfactory bulb
S26
Hippocampus
Cerebellum
Cortex
Olfactory bulb
28S
18S

In situ hybridisation was used to examine in more detail the regional distribution of Hct1 mRNA in the adult brain. A 42-mer oligonucleotide, corresponding to a region of the cDNA 26 bases from the polyA tail (see Figure 4.2(b)) was labelled with ^{35}S-α-dATP using terminal transferase. 10 μm coronal sections of a rat brain, which were cut by Dr. M. Steel, were hybridised with the labelled oligonucleotide. The hybridisation was detected by dipping in photographic emulsion and examining the distribution of silver grains over the sections. Using dark field optics, the distribution of silver grains was visualised as white spots on a dark background (Figure 3.2).

An exposure time of six weeks was required to detect Hct1 expression in the rat brain. The major site of Hct1 expression was in the principal cell layers of the hippocampus (Figure 3.2). Elsewhere in the brain, Hct1 expression was barely detectable above the background signal of hybridisation in the cortex. In the rest of the brain, hybridisation signals equivalent only to the background signal of the silver grains could be detected throughout the rest of the brain sections.

Within the hippocampus, the distribution of Hct1 expression was not uniform (Fig. 3.2(e)). Expression was highest over the cell bodies of the dentate gyrus. Throughout the CA regions of the hippocampus, Hct1 mRNA levels appeared to be graded. Expression was highest in area CA3, but moving through the formation, expression was lower in area CA2, and was barely detectable in the CA1 field.

At higher magnification using bright field optics, the cell bodies of pyramidal neurons within the hippocampus could be visualised, with the silver grains appearing over expressing cells as black spots. Again the expression of Hct1 appeared not to be uniform within sub-structures of the hippocampus (Fig. 3.3). Silver grains were localised to a sub-population of cell bodies within the dentate gyrus, while other cell bodies appeared completely devoid of silver grains. This was also true of the CA3 and CA1 fields of the hippocampus (data not shown). In contrast, an *in situ* hybridisation performed in parallel using a probe to Hct2 (clone 13) (refer to section 4.5) showed silver grains distributed evenly over all the cell bodies of the dentate gyrus (Figure 3.3(b)).

Figure 3.2 Distribution of Hct1 transcripts in the adult rat brain. *In situ* hybridisation using radiolabelled oligonucleotides specific for Hct1 on 10μ coronal sections. Exposure time for autoradiography was six weeks. (Magnification (a -d) x7.5; (e, f) x20).

(a) Bright field view of rat brain coronal section stained with methyl green.
(b) Dark field view of rat coronal section through the hippocampus, hybridised with Hct-1 specific oligonucleotide.
(c) Dark field view of rat coronal section through the hippocampus, hybridised with Hct1-specific oligonucleotide.
(d) Dark field view of rat brain coronal section hybridised with an oligonucleotide derived from opsin (negative control).
(e) Dark field view of hippocampus hybridised with Hct1-specific oligonucleotide..
(f) Bright field view of a coronal section through the hippocampus stained with methyl green.
(Boxed area refers to the region shown in Figure 3.3).

CA1, CA3 pyramidal cell layers; DG, dentate gyrus

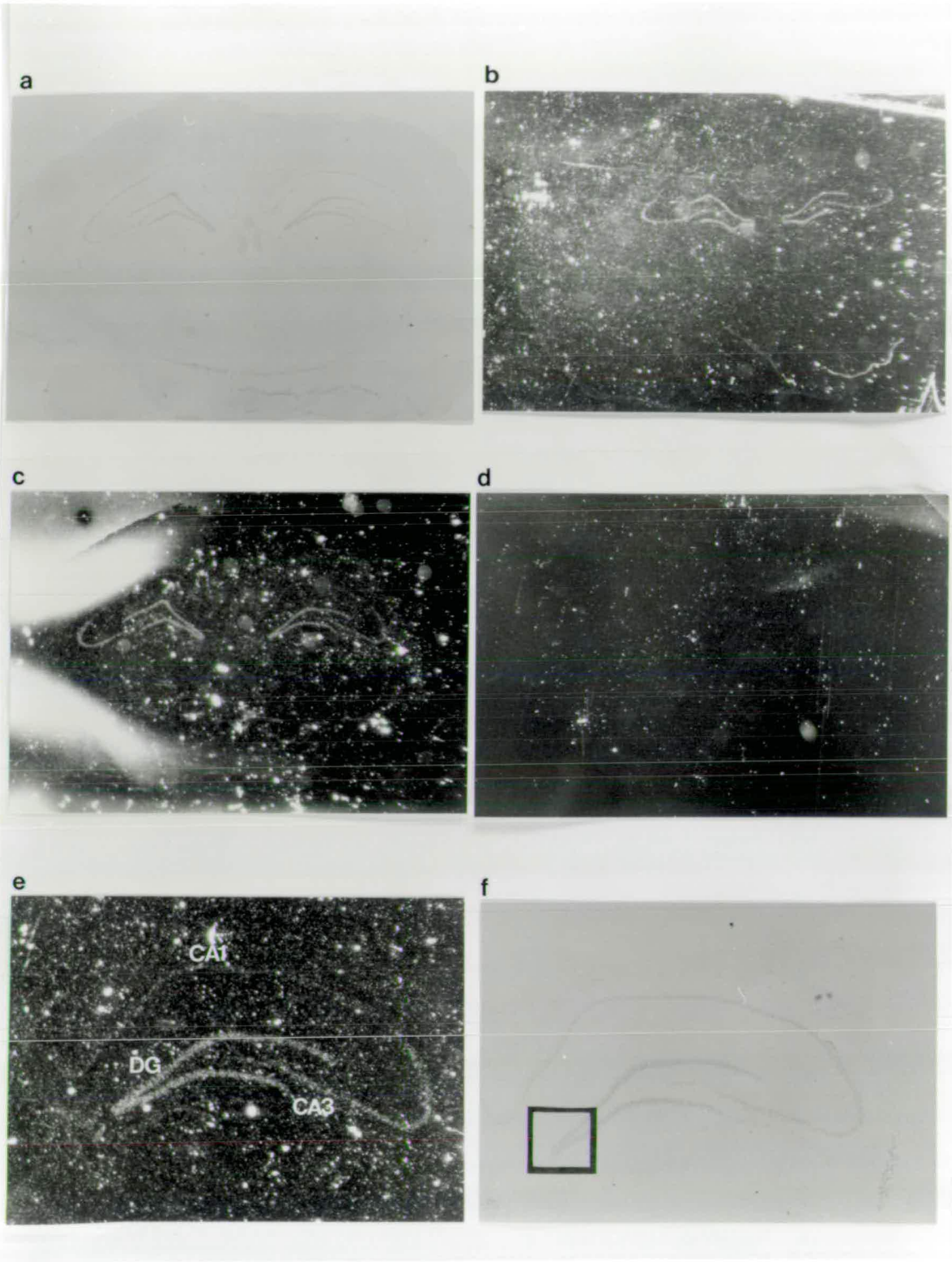
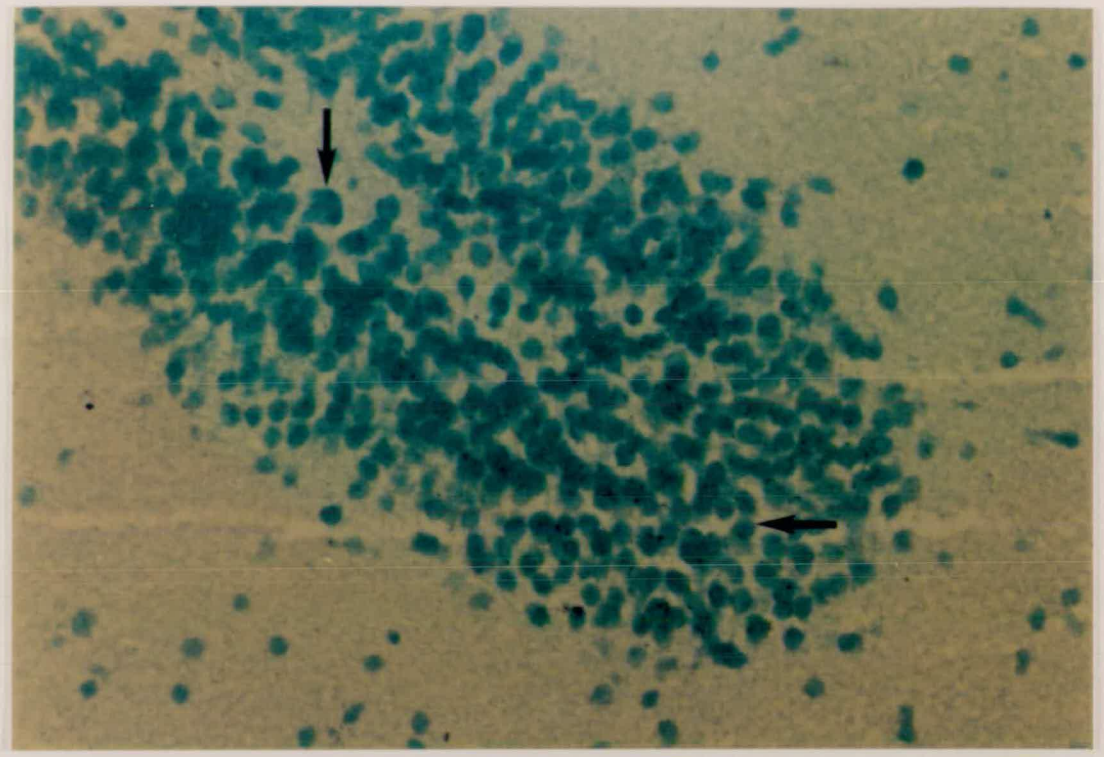

a
b
c
d
e
CA1
DG
CA3
f

Figure 3.3 Cellular localisation of Hct1 transcripts within the hippocampus.

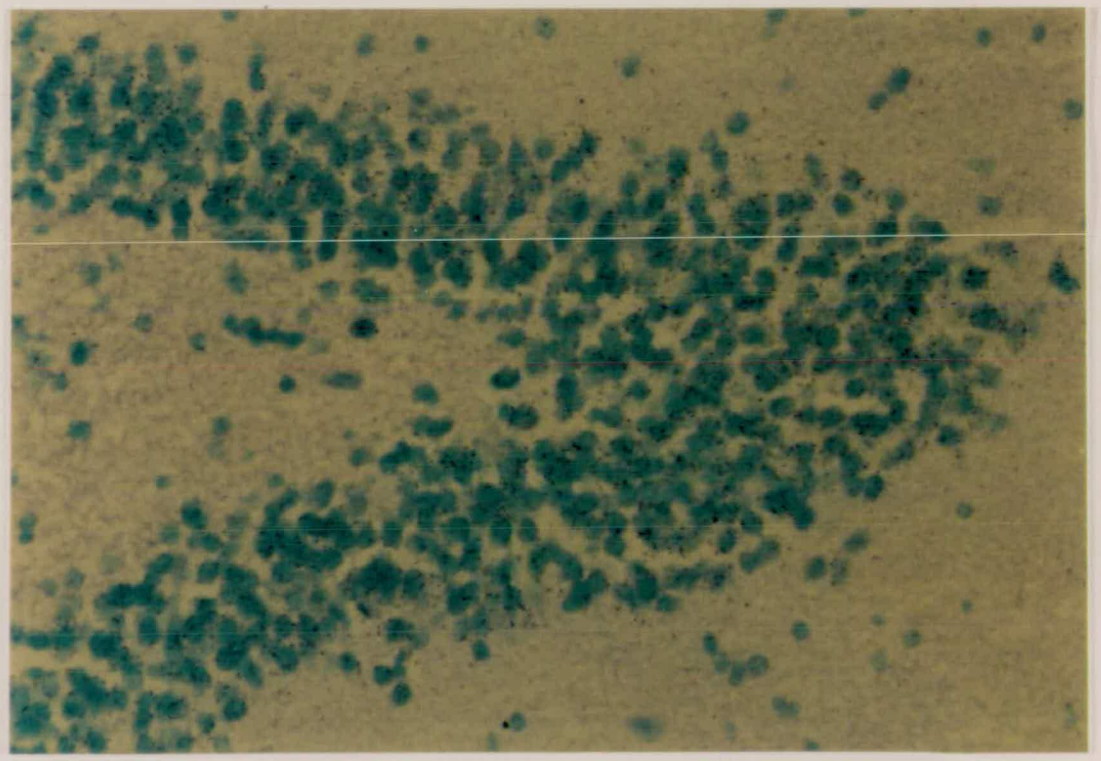
(Figure 3.2(f) indicates the area of the dentate gyrus which is magnified in Figure 3.3)

(a) Bright field high power magnification (x200) of the dentate gyrus granule cells stained with methyl green. The silver grain distribution, showing discrete localisation of Hct1 signal over the granule cells of the dentate gyrus, is not uniform (see arrowed cells).

(b) Hct2 cellular localisation in the dentate gyrus.



(a) Cellular localisation of Hct1 hybridisation signal in rat dentate granule cells.



(b) Cellular localisation of hybridisation signal of Hct2 (clone 13)

Thus it seemed unlikely that the differential distribution of silver grains over the cell bodies was an artefact of the *in situ* hybridisation. Rather, a sub-population of cells in the hippocampus appears to be marked by their expression of Hct1.

3.3 Hct1 mRNA distribution in the adult mouse.

Hct1 mRNA expression was analysed in the mouse. Current technology permits only addition transgenesis to be performed in the rat. Thus conditional ablation technology, as described in Chapter One, could be achieved in the rat to parallel traditional ablation studies already done in the rat. However, homologous recombination in mouse embryonic stem cells would allow for the disruption or modification of the endogenous gene. This technology is not yet developed for the rat, and so the range of genetic manipulations available would be restricted. To expand the range of potential experiments, the expression pattern of Hct1 in the adult mouse brain was also analysed.

The distribution of Hct1 mRNA in mouse was examined by *in situ* hybridisation, however to confirm Hct1 transcript sizes in mouse, a Northern blot of mouse RNA samples was probed with a rat Hct1 probe (Figure 3.4). Northern analysis revealed one transcript of 1.8 kb present in hippocampus, as well as other brain regions and liver. This is in contrast to the three transcripts detected in rat. Expression in kidney was not detected, however a longer exposure may be needed to detect the low level of expression in this tissue

An oligonucleotide based on rat Hct1 sequence was initially used as a probe for the *in situ* hybridisation on mouse brain sections. Later, when a mouse cDNA was cloned for Hct1 (Chapter Four), a mouse 42-mer oligonucleotide was designed to confirm the expression pattern seen with the oligonucleotide derived from rat Hct1 sequence (Figure 3.5).

Figure 3.4 Northern analysis of Hct1 mRNA in mouse tissues.

Total cytoplasmic RNA from mouse tissues was Northern blotted and probed with the radiolabelled rat 1.4 kb cDNA clone 12. Autoradiography was for seven days with intensifying screens. The presence of RNA in all lanes was confirmed by reprobing the filter with a probe for ribosomal protein S26.

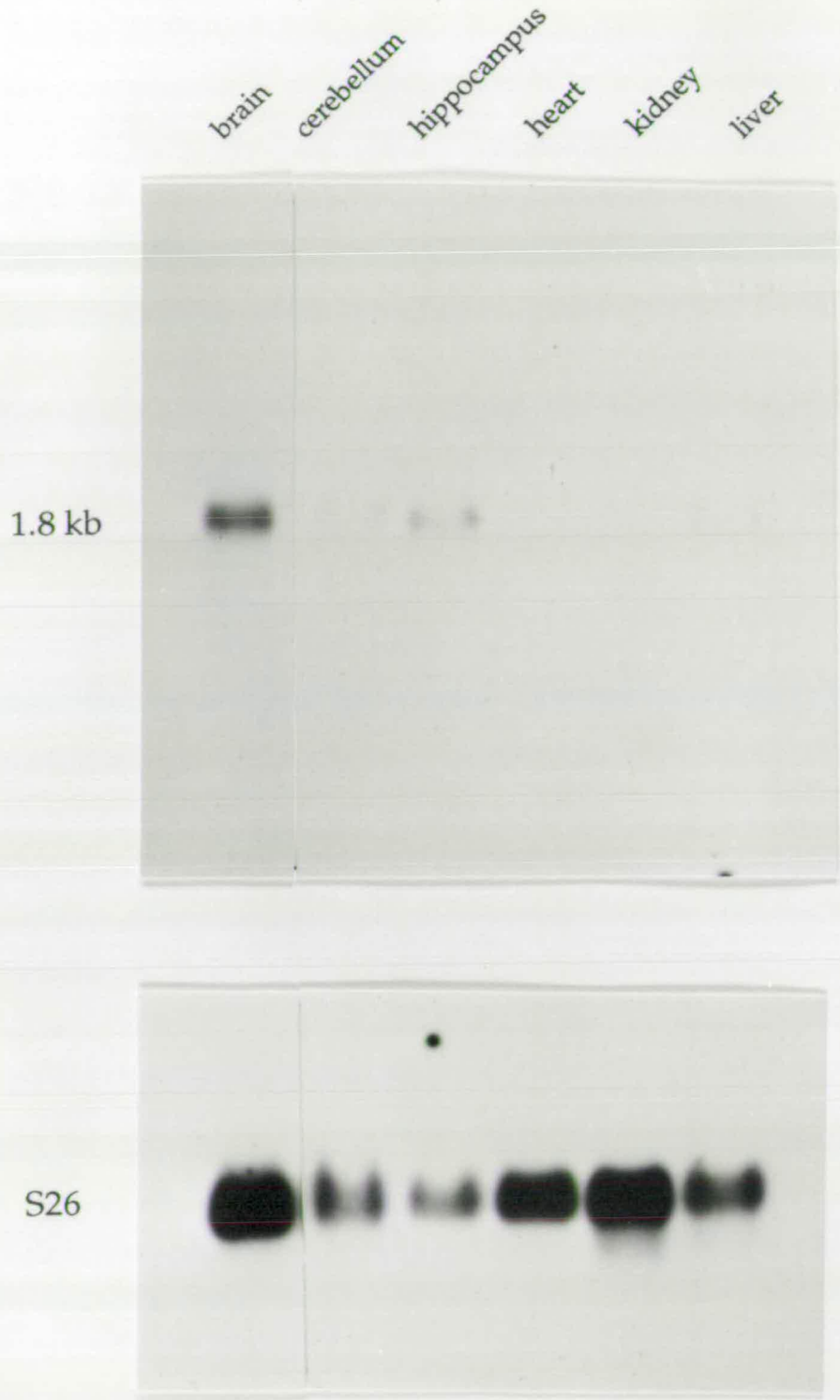
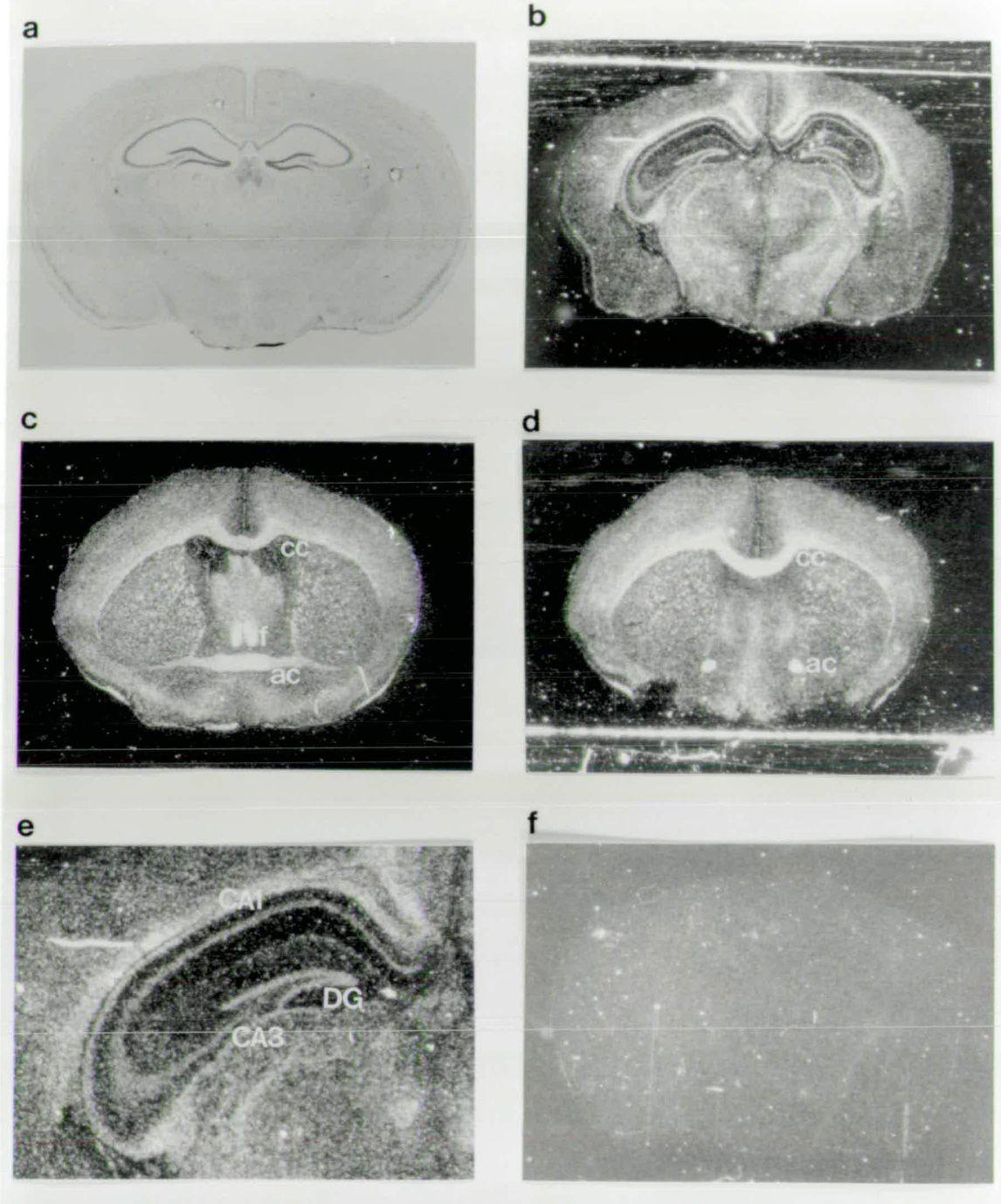

brain
cerebellum
hippocampus
heart
kidney
liver
1.8 kb
S26

Figure 3.6 *In situ* hybridisation analysis of Hct1 in the adult mouse brain.

(x7.5 magnification)
(a) Bright field view of mouse coronal section through the hippocampus, stained with 1% neutral red (section provided by Dr. M. Steel).
(b) Dark field view of mouse coronal section through the hippocampus, hybridised with a Hct1-specific oligonucleotide.
(c) Dark field view of mouse coronal section (rostral to section in (b)), showing strong Hct1 signal in corpus callosum, fornix and anterior commissure.
(d) Dark field view of mouse brain coronal section (rostral to section in (c)), showing strong Hct1 signal in corpus callosum and anterior commissure.
(e) Dark field view of Hct1 signal in a coronal section through the hippocampus (x20 magnification).
(f) Dark field view of mouse brain coronal section hybridised with an oligonucleotide specific for opsin (negative control).

CA1, CA3 pyramidal cell layers; DG, dentate gyrus; cc, corpus callosum; ac anterior commissure; f, fornix


a
b
c
cc
ac
d
cc
ac
e
CA1
DG
CA3
f

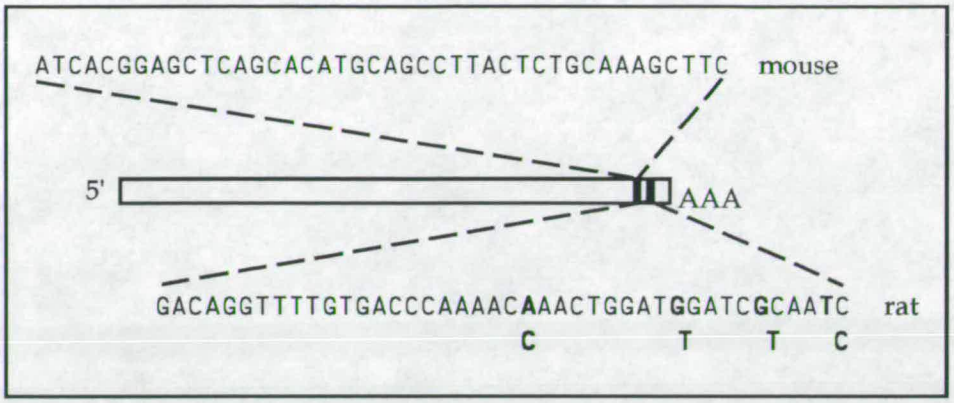


Figure 3.5 Relative positions of oligonucleotides used for Hct1 *in situ* hybridisation in mouse and rat.
The rat oligonucleotide was situated 37 nucleotides from the beginning of the polyA tail. The four nucleotide differences between mouse and rat have been indicated. The mouse oligonucleotide was situated a further 58 nucleotides upstream from the position of the rat oligonucleotide.

Hct1 was found to be widely expressed throughout the adult mouse brain (Figure 3.6). In the hippocampus, expression was highest in the dentate gyrus as was found in rat. However in contrast to rat, Hct1 mRNA was detected at equivalent levels in areas CA3, CA2 and CA1. In addition, within the hippocampal formation, a layer of expression was detected outside of the principal cell bodies which may correspond either to interneuron expression or dendritic mRNA localisation (Figure 3.6(e)). Elsewhere in the mouse brain, expression of Hct1 was widespread. The strongest expression of Hct1 in the mouse brain was detected in the corpus callosum, the anterior commissure and fornix (Figure 3.6(c) (d)). Moderate expression levels, comparable to that seen in the hippocampus, were also observed in the cerebellum, cortex and olfactory bulb (data not shown).

3.4 Discussion

Analysis of the expression pattern of Hct1 in the rat has revealed a distinct distribution of Hct1 transcripts. Extra-neuronal tissues expressing Hct1, the liver and kidney, contain very low levels of Hct1

relative to the brain. *In situ* hybridisation revealed not only the hippocampus as the predominant site of expression, but also a distinct pattern of expression within the hippocampus itself. A subset of neurons within the principal cell layers of the hippocampus express Hct1 while adjacent cells do not. The cell bodies within this layer are predominantly the excitatory pyramidal cells in the CA areas and the granule cells in the dentate gyrus. However, the cell layers are not homogeneous, and interneurons such as the three different types of basket cells are also present in these layers. Hct1, therefore, might be produced only by a particular cell type within these layers. Alternatively, Hct1 expression may be highlighting differences between morphologically indistinguishable pyramidal or granule cells within the hippocampus. Double *in situ* labelling techniques at either the RNA or protein level could address this question. For example, colocalisation of Hct1 with other known cell markers such as GABA (γ-amino-butyric acid), a marker for GABAergic inhibitory interneurons, could confirm Hct1 expression as confined to a defined subset of cells within the hippocampus. Nevertheless, such a specific pattern of expression for Hct1 suggests a defined role for this gene product within the hippocampus. With the use of antibodies it would be possible to address Hct1 distribution at an even higher level of resolution, that is, the subcellular distribution of Hct1 within the neuron. For example, localised gene expression only within the dendritic spines of a neuron could support a role for the gene product at the post-synaptic cell. Expression of the protein kinase, CaMKII, for example, is confined to the post-synaptic density (Schulman and Lou, 1989)

Expression of mRNA in a sub-population of cells within the hippocampal formation has been noted in other *in situ* analyses. mRNA for brain-derived neurotrophic factor (BDNF) has been shown to be confined to a subset of neurons in all regions of the hippocampus (Hofer *et al.*, 1990). However, while the authors suggested differential expression within the principal pyramidal neurons of the different CA regions and granule cells of the dentate gyrus, equally the differential expression could be due to the cell bodies of a discrete type of neuron such as interneurons which are also dispersed in this layer of cells.

Nerve growth factor (NGF), too, has been shown to be differentially expressed within cells of the hippocampus (Ayer-Lelievre *et al.*, 1988).

The use of Northern and *in situ* hybridisation to analyse rat Hct1 expression has yielded different results in terms of relative distribution of the transcript within the brain. Northern analysis gave no indication of enrichment of Hct1 expression in rat hippocampus, while *in situ* hybridisation indicated preferential expression in this tissue.

The use of ribosomal protein S26 mRNA levels to standardise sample loading is open to question. Ethidium bromide staining of rRNA bands to standardise sample loading is not optimal as it is not sensitive enough to detect subtle differences in loading. In addition, this method does not control for relative RNA amounts after the RNA is transferred from the gel to the membrane. It must therefore be assumed that the efficiency of transfer was equivalent for all samples. Using this latter method, S26 levels were shown to vary in the hippocampus which affected interpretation of Hct1 expression within different brain regions. While neither standardisation method proved satisfactory, the Hct1 expression pattern was decided by *in situ* hybridisation.

When sample loading was standardised according to relative ethidium bromide staining of rRNA, the Northern data still did not reflect the *in situ* results. The discrepancies observed may perhaps be explained from the *in situ* results, where only a subset of hippocampal neurons is expressing Hct1. Thus, the prepared hippocampus RNA is diluted by a proportion of non-expressing cells. While in the cortex or olfactory bulb, for example, although expression of Hct1 is very low, all cells may be expressing it (This could be confirmed by counting silver grains in these brain regions, however the level is barely above background hybridisation and would thus be difficult to interpret.) Thus RNA prepared from these two tissues and then analysed by Northern blot would mask the relative amounts of Hct1 expressed in each sample. This dilution effect can also be seen in Figure 3.1(a), where the Hct1 signal from "remainder of brain" relative to cerebellum, cortex and olfactory bulb samples is much lower, even though these samples are represented in "ROB".

From this study, it would seem that *in situ* hybridisation offers the more accurate method of detecting relative expression levels, while Northern analysis provides a primary level of detection and transcript patterns between different tissues. This also suggests that the use of Northern analysis to identify differential Hct clones was not perhaps the method of choice, and that the use of *in situ* hybridisation to identify potential Hct clones may be more informative.

A striking difference in Hct1 expression patterns was observed between the rat and mouse. While both species express Hct1 in the liver at similar levels, the distribution of Hct1 throughout the adult brain is quite different. Rat Hct1 is highly enriched in the hippocampus, with very low levels detected elsewhere in the brain by *in situ* hybridisation. This is consistent with the differential screen by which it was isolated. However, mouse Hct1 is not enriched in the hippocampus, but appears to be most abundant in the corpus callosum. Further, transcripts of mouse Hct1 are widespread throughout the brain, and are present at much higher levels than in rat. Mouse Hct1 transcripts could be detected by *in situ* hybridisation after a three week exposure, while in parallel experiments with the same probe, detection of the rat orthologue required exposure times of six to eight weeks.

Within the hippocampus, the expression of Hct1 in mouse and rat is not conserved. While both express Hct1 to similar levels in the dentate gyrus, expression in the CA areas of the hippocampus are not conserved. Rat Hct1 is expressed in area CA3, with barely detectable levels in CA1, while mouse Hct1 is expressed equally throughout the CA fields of the hippocampus.

The possibility that the oligonucleotide used in the mouse experiments cross-reacted with a related gene cannot be excluded but seems unlikely as the oligonucleotides correspond to the 3' ends of the cDNAs, a region which tends to be less conserved between paralogues. In addition, both rat and mouse oligonucleotides on mouse brain sections gave the same pattern of expression. Thus, unless the related gene is present in mouse but not in rat, the *in situ* hybridisation results have revealed an example of non-conserved orthologous gene expression.

Other examples of non-conserved expression patterns of orthologous genes are rare in the literature. This may be in part because detailed *in situ* expression patterns of genes are rarely checked in different species. The Il-1 receptor, however, appears to be widely distributed in the rat brain, with highest densities in the cerebellar cortex, olfactory bulb and choroid plexus. The mouse expression pattern is more restricted, with highest expression in the granule cell layer of the dentate gyrus and the choroid plexus (Linthorst *et al.*, 1994 and references therein). A serotonin receptor, 5-HT$_{5B}$, which is expressed at very high levels in the rat hippocampus CA1 cells with low levels detected in the medial habenulae (Wisden *et al.*, 1993), has a reversed expression pattern in the mouse (Wisden, personal communication). Expression of mouse 5-HT$_{5B}$ is very high in the habenulae, but very low in the hippocampus .

While the non-conservation of the expression pattern for Hct1 does not support a conserved or required function for the gene product, Hct1 in the rat is expressed in a spatially regulated fashion, resulting in significant hippocampus enrichment. The expression pattern of Hct1 in rat prompted further cloning of Hct1 and characterisation of the gene product.

CHAPTER FOUR

GENOMIC AND cDNA STRUCTURE OF CLONE 14.5a (Hct1)

4.1 Introduction

Hct1 gene expression in the rat was shown to be highly enriched in the hippocampus of the adult brain. However, because the original clone of Hct1, 14.5a, was only 0.3 kb in length, it yielded very little information about the gene product. Molecular cloning of Hct1 transcripts was used to provide more information about the gene product of Hct1. Once the nucleotide sequence of a full-length cDNA is obtained, the amino acid sequence of the gene product can be deduced, and compared with databases of known sequences for homologies and the presence of conserved motifs.

Hct1 appears to have a complex pattern of transcription, as three mRNAs of 1.8 kb, 2.2 kb, and 5.0 kb were detected by Northern analysis. Variable transcripts from a transcription unit may be formed by a number of different mechanisms, including alternative mRNA splicing, the use of different promoters, and differential polyadenylation within the 3' untranslated region of the gene (for review, see McKeown, 1992). Such transcript variability would provide a level of regulation, where different forms of a gene product might be produced, each with possibly different functions. An interesting example of new gene functions as a result of splicing is described by Sommer *et al.* (1990), where the inclusion of alternatively spliced exons (flip and flop) alters the electrophysiological properties of the AMPA-selective glutamate-gated ion channels. To investigate the transcriptional complexity of Hct1, rat cDNA libraries were screened with clone 14.5a as a probe in order to isolate further Hct1 cDNAs.

Mouse cDNA libraries and mouse genomic DNA libraries were also screened to isolate the mouse orthologue and to investigate the

organisation of the Hct1 gene, with a view to the identification of transcriptional regulatory elements controlling Hct1 expression.

4.2 Hct1 cDNA structure

The rat hippocampus cDNA library used in the initial differential hybridisation screen was replated so that 10^6 recombinant *lambda* ZAP II phage plaques were screened at a density of 50,000 per 15 cm plate. Duplicate filter lifts were made from each plate, and these were hybridised to clone 14.5a, labelled by random-primer labelling. A total of 21 mixed primary clones were isolated and subjected to secondary screening to confirm the clones and to purify each. Of these, only four isolates gave a strong signal on rescreening, and these were further purified.

pBluescript II SK plasmids were rescued from each purified *lambda* ZAP II clone by *in vivo* excision, and the insert sizes were determined using EcoRI and NotI digests. To confirm that these clones were related to 14.5a, the 3' sequence adjacent to the polyA tail was determined using primers which hybridised to the T3 and T7 RNA polymerase sites flanking the cloning site in the pBluescript II SK vector (Figure 4.1).

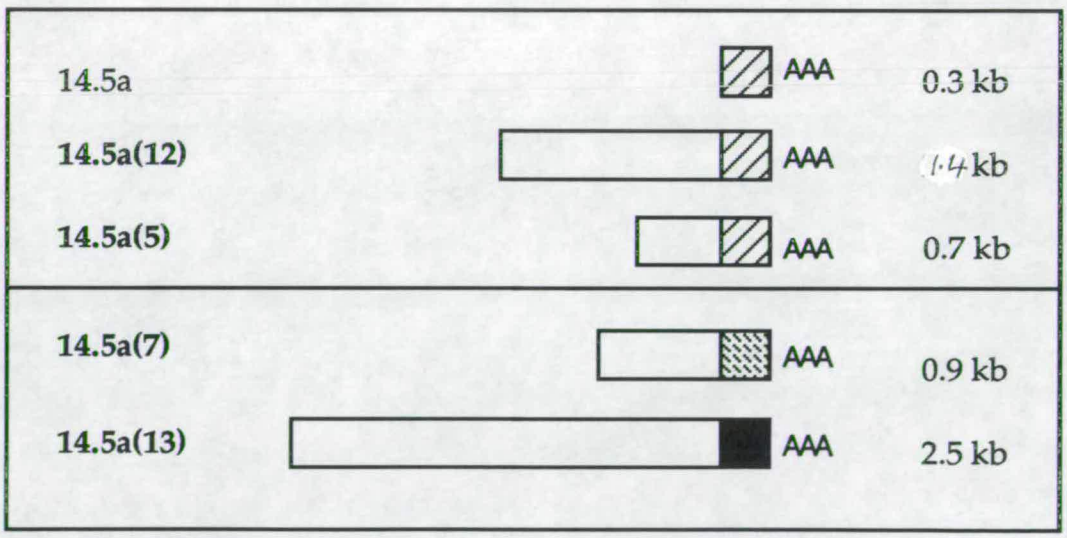


Figure 4.1 cDNA clones isolated from a rat hippocampus cDNA library using an Hct1 probe. The sizes of each were determined by EcoRI and NotI digests of the recombinant plasmids. The hatched boxes indicate those clones with the same 3' end as clone 14.5a. Filled and broken hatched boxes represent different 3' sequences.

4.2(a) The 1.8 kb transcript

Clones 5 and 12 were shown to have the same 3' end as the original clone, while the 3' sequences of clone 7 and clone 13 were different to the original 14.5a clone and different to each other.

Clone 12, being the largest clone isolated with 3' sequence homology to clone 14.5a (1.4 kb), was chosen for further sequencing and restriction mapping. Exonuclease III deletions were created in the cDNA clone and these deletions were sequenced directly using a primer which hybridised to the T7 polymerase binding site. Sequence contigs of one strand were assembled to generate the complete sequence of clone 12 (Figure 4.2).

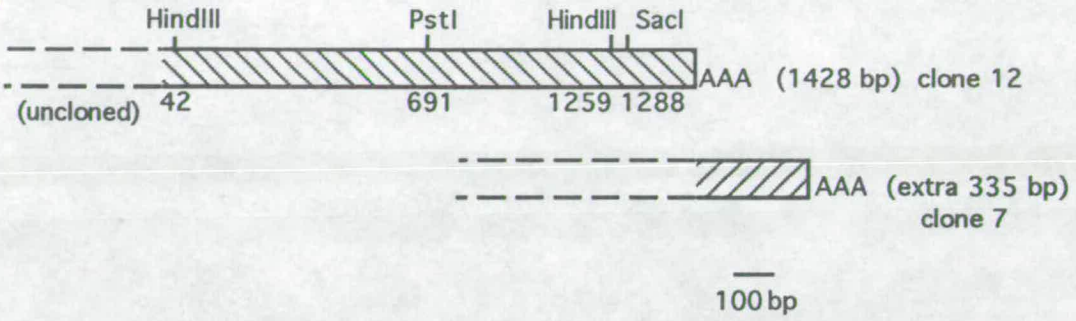
The 1428 bp sequence of clone 12 was compared with previously described sequences in the GenBank/EMBL sequence data base. Significant homology between clone 12 and the human and rat cDNAs for cholesterol 7α-hydroxylase (Noshiro and Okuda, 1990; Li *et al*, 1990) was found. At the nucleic acid level, the 1428 bp cDNA clone for rat Hct1 shares 54.6% identity over a 1100 bp overlap with human cholesterol 7α-hydroxylase (*CYP7*) and 53.6% identity over a 1117 bp overlap with the rat *CYP7* cDNA. Translated sequence from rat Hct1 shared 37.2 and 36.7% amino acid identity with the human and rat CYP7 proteins respectively. The 5' end of clone 5 was sequenced and was found to be contained within the 1.4 kb sequence, its 0.7 kb size correlating with the position of homology within clone 12. Clone 5 was therefore a shorter cDNA of clone 12.

The Hct1 transcripts observed by Northern analysis (1.8 kb, 2.2 kb and 5.0 kb) are all larger than clone 12, and in addition, the homology to cholesterol 7α-hydroxylase was truncated at the 5' end. Clone 12 is at least 400 bp short of the smallest full length transcript.

Figure 4.2 a) restriction map of clone 12.
Restriction enzymes which cut within the pBluescript polylinker were used to map the restriction sites of clone 12. The only such enzymes that cut the insert were HindIII at nucleotides 42 and 1259, PstI at 691, and SacI at 1288.

b) complete nucleotide and translation sequence of the 1.4 kb cDNA clone of rat Hct1 (clone 12), including additional clone 7 sequence. Clone 12 is 1428 bp (upper case), with an open reading frame from nucleotide 1 to a TGA stop codon at 1243. The additional 335 bp of 3' sequence corresponding to the novel sequence of clone 7 is indicated by lower case sequence. The two putative polyadenylation signals are underlined. (The conserved cysteine referred to in Chapter Five is boxed).

(a)



(b)

```
A  L  E  Y  Q  Y  V  M  K  N  P  K  Q  L  S  F  E  K  F  S  R  R
GCCTTGGAGTACCAGTATGTAATGAAAAACCCAAAACAATTAAGCTTTGAGAAGTTCAGCCGAAG    65

 L  S  A  K  A  F  S  V  K  K  L  L  T  N  D  D  L  S  N  D  I  H
ATTATCAGCGAAAGCCTTCTCTGTCAAGAAGCTGCTAACTAATGACGACCTTAGCAATGACATTC   130

  R  G  Y  L  L  L  Q  G  K  S  L  D  G  L  L  E  T  M  I  Q  E
ACAGAGGCTATCTTCTTTTACAAGGCAAATCTCTGGATGGTCTTCTGGAAACCATGATCCAAGAA   195

V  K  E  I  F  E  S  R  L  L  K  L  T  D  W  N  T  A  R  V  F  D
GTAAAAGAAATATTTGAGTCCAGACTGCTAAAACTCACAGATTGGAATACAGCAAGAGTATTTGA   260

 F  C  S  S  L  V  F  E  I  T  F  T  T  I  Y  G  K  I  L  A  A  N
TTTCTGTAGTTCACTGGTATTTGAAATCACATTTACAACTATATATGGAAAAATTCTTGCTGCTA   325

  K  K  Q  I  I  S  E  L  R  D  D  F  L  K  F  D  D  H  F  P  Y
ACAAAAAACAAATTATCAGTGAGCTGAGGGATGATTTTTTAAAATTTGATGACCATTTCCCATAC   390

L  V  S  D  I  P  I  Q  L  L  R  N  A  E  F  M  Q  K  K  I  I  K
TTAGTATCTGACATACCTATTCAGCTTCTAAGAAATGCAGAATTTATGCAGAAGAAAATTATAAA   455

 C  L  T  P  E  K  V  A  Q  M  Q  R  R  S  E  I  V  Q  E  R  Q  E
ATGTCTCACACCAGAAAAAGTAGCTCAGATGCAAAGACGGTCAGAAATTGTTCAGGAGAGGCAGG   520

  M  L  K  K  Y  Y  G  H  E  E  F  E  I  G  A  H  H  L  G  L  L
AGATGCTGAAAAAATACTACGGGCATGAAGAGTTTGAAATAGGAGCACATCATCTTGGCTTGCTC   585

W  A  S  L  A  N  T  I  P  A  M  F  W  A  M  Y  Y  L  L  Q  H  P
TGGGCCTCTCTAGCAAACACCATTCCAGCTATGTTCTGGGCAATGTATTATCTTCTTCAGCATCC   650

 E  A  M  E  V  L  R  D  E  I  D  S  F  L  Q  S  T  G  Q  K  K  G
AGAAGCTATGGAAGTCCTGCGTGACGAAATTGACAGCTTCCTGCAGTCAACAGGTCAAAAGAAAG   715

  P  G  I  S  V  H  F  T  R  E  Q  L  D  S  L  V  C  L  E  S  A
GACCTGGAATTTCTGTCCACTTCACCAGAGAACAATTGGACAGCTTGCTCTGCCTGGAAAGCGCT   780
```

```
                                                                    .
I  L  E  V  L  R  L  C  S  Y  S  S  I  I  R  E  V  Q  E  D  M  D
ATTCTTGAGGTTCTGAGGTTGTGCTCCTACTCCAGCATCATCCGTGAAGTGCAAGAGGATATGGA   845
   .     .     .     .     .     .     .     .     .     .     .

 F  S  S  E  S  R  S  Y  R  L  R  K  G  D  F  V  A  V  F  P  P  M
TTTCAGCTCAGAGAGTAGGAGCTACCGTCTGCGGAAAGGAGACTTTGTAGCTGTCTTTCCTCCAA   910
      .     .     .     .     .     .     .     .     .     .

   I  H  N  D  P  E  V  F  D  A  P  K  D  F  R  F  D  R  F  V  E
TGATACACAATGACCCAGAAGTCTTCGATGCTCCAAAGGACTTTAGGTTTGATCGCTTCGTAGAA   975
   .     .     .     .     .     .     .     .     .     .     .

D  G  K  K  K  T  T  F  F  K  G  G  K  K  L  K  S  Y  I  I  P  F
GATGGTAAGAAGAAAACAACGTTTTTCAAAGGAGGAAAAAAGCTGAAGAGTTACATTATACCATT   1040
      .     .     .     .     .     .     .     .     .     .

 G  L  G  T  S  K  C  P  G  R  Y  F  A  I  N  E  M  K  L  L  V  I
TGGACTTGGAACAAGCAAATGTCCAGGCAGATACTTTGCAATTAATGAAATGAAGCTACTAGTGA   1105
   .     .     .     .     .     .     .     .     .     .     .

   I  L  L  T  Y  F  D  L  E  V  I  D  T  K  P  I  G  L  N  H  S
TTATACTTTTAACTTATTTTGATTTAGAAGTCATTGACACTAAGCCTATAGGACTAAACCACAGT   1170
      .     .     .     .     .     .     .     .     .     .

R  M  F  L  G  I  Q  H  P  D  S  D  I  S  F  R  Y  K  A  K  S  W
CGCATGTTTCTGGGCATTCAGCATCCAGACTCTGACATCTCATTTAGGTACAAGGCAAAATCTTG   1235
   .     .     .     .     .     .     .     .     .     .     .

 R  S  •••
GAGATCCTGAAAGGGTGGCAGAGAAGCTTAGCGGAATAAGGCTGCACATGCTGAGCTCTGTGATT   1300
.     .     .     .     .     .     .     .     .     .     .
TGCTGTACTCCCCAAATGCAGCCACTATTCTTGTTTGTTAGAAAATGGCAAATTTTTATTTGATT   1365
.     .     .     .     .     .     .     .     .     .     .
GCGATCCATCCAGTTTGTTTTGGGTCACAAAACCTGTCATAAAATAAAGCGCTGTCATGGTGTaa   1430
.     .     .     .     .     .     .     .     .     .     .
aaaaatgtcatggcaatcatttcaggataaggtaaaataacgttttcaagtttgtacttactatg   1495
.     .     .     .     .     .     .     .     .     .     .
atttttatcatttgtagtgaatgtgcttttccagtaataaatttgcgccagggtgatttttttta   1560
.     .     .     .     .     .     .     .     .     .     .
attactgaaatcctctaatatcggttttatgtgctgccagaaaagtgtgccatcaatggacagta   1625
.     .     .     .     .     .     .     .     .     .     .
taacaatttccagttttccagagaagggagaaattaagccccatgagttacgctgtataaaattg   1690
.     .     .     .     .     .     .     .     .     .     .
ttctcttcaactataatatcaataatgtctatatcaccaggttacctttgcattaaatcgagttt   1755
.     .
tgcaaaag   1763
```

4.2(b) The 2.2 kb transcript

DNA sequencing of clone 7 revealed that its 3' end is unrelated to the 3' sequence of the original 300 bp probe or of clone 12. Restriction mapping of clone 7 did however reveal the same juxtaposed SacI and HindIII sites found in clone 12 (Figure 4.2(a)) (data not shown). In addition, sequencing of the 5' end of clone 7 showed that this sequence was contained within clone 12. Deletions were made in clone 7 using exonuclease III and the subclones were sequenced and assembled into contigs (Figure 4.2). Clone 7 and clone 12 share 594 bp of sequence up to the 3' end of clone 12, however 335 bp of novel 3' sequence, ending in a polyA tail, forms the unique 3' sequence of clone 7. From this analysis, it seems that clone 7 is derived from the same transcription unit and is generated through the use of an alternative polyadenylation site. A potential polyadenylation signal, 5'-AUUAAA-3', was found 22 bases upstream of the polyA tail of clone 7 (Fig. 4.2(b)). The extra 335 bp of the 3' end of clone 7 is also in agreement with the observed 2.2 kb transcript by Northern analysis.

4.3 Cloning of the Hct1 mouse orthologue (mHct1).

The longest rat cDNA clone for Hct1 (1.4 kb) is approximately 400 bp shorter than the expected full length cDNA, based on the sequence homology to cholesterol 7α-hydroxylase and the observed minimum transcript size (see Figure 3.1). With a view to possible future transgenic experiments, particularly gene-targeting, it was decided to screen mouse cDNA libraries for a full length cDNA clone of Hct1. Northern analysis of mouse tissues (refer to Chapter 3, Figure 3.5) using the rat 1.4 kb cDNA as a probe confirmed the existence of Hct1 in the mouse.

A mouse liver cDNA library (a gift from Dr. B. Luckow, Heidelberg) was screened for Hct1 cDNAs. 10^6 recombinant *lambda* gt10 plaques were plated at a density of 50,000 pfu/15 cm plate, duplicate lifts were prepared, and each was probed with the 1.4 kb rat Hct1 clone. Approximately 400 duplicate plaques gave positive signals, and fifteen

clones were purified after a secondary screen. The inserts from each *lambda* gt10 clone were excised using EcoRI and NotI restriction enzymes. These inserts were subcloned into a pBluescript II KS vector.

Nine of eleven clones picked contained the largest insert size of approximately 1.8 kb. DNA sequencing demonstrated that three clones, 33, 35 and 38 were identical at their 5' and 3' ends, and that this sequence was 71% similar to the 3' end of rat Hct1. Thus the mouse clones appeared to encode Hct1. Clones 23 and 40 were also identical to each other, and were related to the first three clones except for 59 bp extra sequence at the 5' end, and 100 bp less at the 3' end. Clones 35 and 40 were chosen for further characterisation. Exonuclease III deletions were made in the cDNA inserts. These deleted subclones were sequenced, and the contigs assembled to produce a full length sequence of the clones (Figure 4.3(b)).

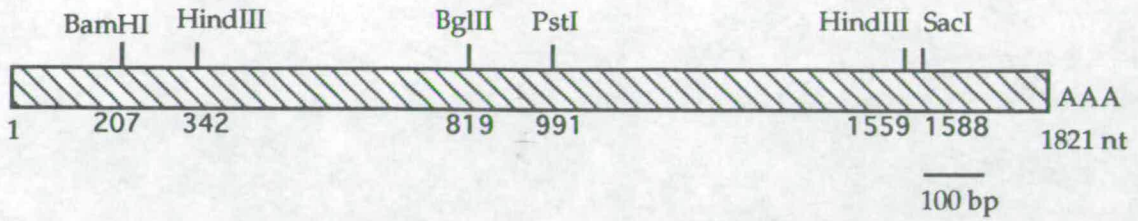
The mouse Hct1 open reading frame commences with a methionine encoded at nucleotide number 22, and terminates with a TGA codon at nucleotide number 1543, encoding a protein of 507 amino acids. The cDNA clone 40 for mHct1 contains an extra 59 nucleotides at the 5' end, however none of this sequence encoded an in frame methionine. cDNA clone 40 is also truncated at the 3' end, being 100 nucleotides shorter than clone 35, without affecting the open reading frame. No polyA tail was associated with either clone and no appropriately positioned polyadenylation signal can be found. Thus some 3' untranlated sequence may be missing from the clone. However the size of the clone and the size of the transcript by Northern analysis predict that very little 3' sequence is missing.

The consensus Kozak sequence for the initiation of translation of a messenger RNA is YYAYYATGR (Kozak, 1987). There are two ATG codons in the mHct1 cDNA, only the first being in frame, however neither sequence conforms to the consensus Kozak sequence. Point mutation analysis of the sequence surrounding the initiator ATG have shown the importance of purines three nucleotides upstream of the initiator codon and immediately 3' of the ATG (Kozak, 1986). Because neither ATG in Hct1 appears to conform to the expected sequence, it

Figure 4.3 (a) Restriction map of mouse Hct1 cDNA.
A restriction map of mHct1 (clone 35) was constructed using restriction enzymes such as those found in the polylinker of pBluescript. Restriction enzymes found to cut mHct1 were BamHI at nucleotide 207, HindIII at 342 and 1559, BglII at 819, PstI at 991, and SacI at 1588.

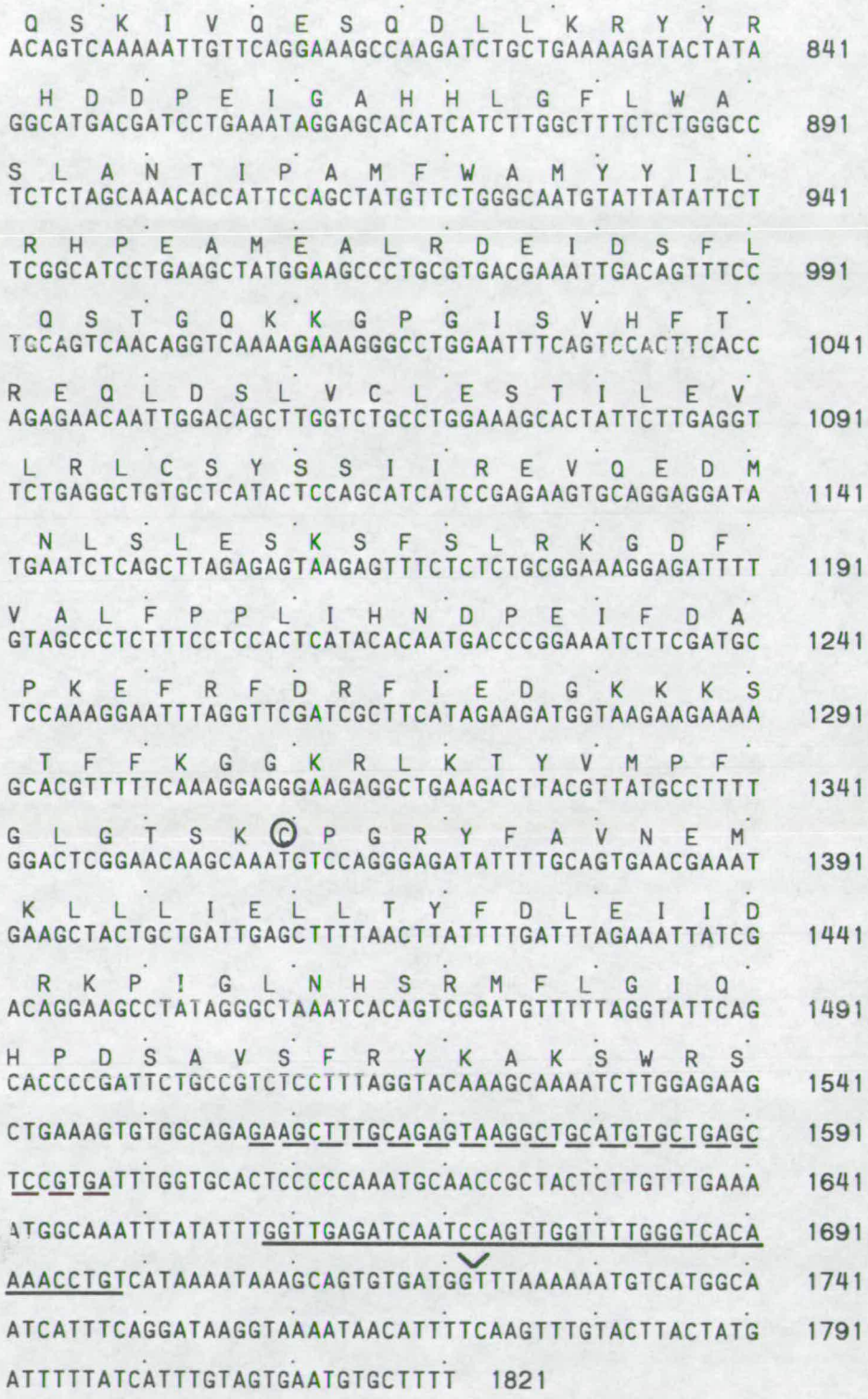
(b) Nucleotide sequence and translation of mHct1. mouse Hct1 refers to clone 35, and all numbering is based on clone 35. Lower case sequence indicates the 59 extra nucleotides in clone 40; the arrow head indicates the end of clone 40. Underlined sequence ( ) represents the oligonucleotide used for *in situ* hybridisation (see section 3.2), and (- -) marks the relative position of the rat oligonucleotide for *in situ* hybridisation. The conserved cysteine (amino acid 447) in the haem-binding domain (see Chapter 5) is circled.

(a)



(b)

```
ggcaggcacagcctctggtctaagaagagagggcactgtgcagaagccat
                                  M  Q  G  A  T  T  L
cgctccctaCAGAGCCGCCAGCTCGTCGGGATGCAGGGAGCCACGACCCT    41
 D  A  A  S  P  G  P  L  A  L  L  G  L  L  F  A  A
AGATGCCGCCTCGCCAGGGCCTCTCGCCCTCCTAGGCCTTCTCTTTGCCG    91
 T  L  L  L  S  A  L  F  L  L  T  R  R  T  R  R
CCACCTTACTGCTCTCGGCCCTGTTCCTCCTCACCCGGCGCACCAGGCGC   141
 P  R  E  P  P  L  I  K  G  W  L  P  Y  L  G  M  A
CCTCGTGAACCACCCTTGATAAAAGGTTGGCTTCCTTATCTTGGCATGGC   191
 L  K  F  F  K  D  P  L  T  F  L  K  T  L  Q  R  Q
CCTGAAATTCTTTAAGGATCCGTTAACTTTCTTGAAAACTCTTCAAAGGC   241
 H  G  D  T  F  T  V  F  L  V  G  K  Y  I  T  F
AACATGGTGACACTTTCACTGTCTTCCTTGTGGGGAAGTATATAACATTT   291
 V  L  N  P  F  Q  Y  Q  Y  V  T  K  N  P  K  Q  L
GTTCTGAACCCTTTCCAGTACCAGTATGTAACGAAAAACCCAAAACAATT   341
 S  F  Q  K  F  S  S  R  L  S  A  K  A  F  S  V  K
AAGCTTTCAGAAGTTCAGCAGCCGATTATCAGCGAAAGCCTTCTCTGTAA   391
 K  L  L  T  D  D  D  L  N  E  D  V  H  R  A  Y
AGAAGCTGCTTACTGATGACGACCTTAATGAAGACGTTCACAGAGCCTAT   441
 L  L  L  Q  G  K  P  L  D  A  L  L  E  T  M  I  Q
CTACTTCTACAAGGCAAACCTTTGGATGCTCTTCTGGAAACTATGATCCA   491
 E  V  K  E  L  F  E  S  Q  L  L  K  I  T  D  W  N
AGAAGTAAAAGAATTATTTGAGTCCCAACTGCTAAAAATCACAGATTGGA   541
 T  E  R  I  F  A  F  C  G  S  L  V  F  E  I  T
ACACAGAAAGAATATTTGCATTCTGTGGCTCACTGGTATTTGAGATCACA   591
 F  A  T  L  Y  G  K  I  L  A  G  N  K  K  Q  I  I
TTTGCGACTCTATATGGAAAAATTCTTGCTGGTAACAAGAAACAAATTAT   641
 S  E  L  R  D  D  F  F  K  F  D  D  M  F  P  Y  L
CAGTGAGCTAAGGGATGATTTTTTTAAATTTGATGACATGTTCCCATACT   691
 V  S  D  I  P  I  Q  L  L  R  N  E  E  S  M  Q
TAGTATCTGACATACCTATTCAGCTTCTAAGAAATGAAGAATCTATGCAG   741
K  K  I  I  K  C  L  T  S  E  K  V  A  Q  M  Q  G
AAGAAAATTATAAAATGCCTCACATCAGAAAAAGTAGCTCAGATGCAAGG   791
```


Q S K I V Q E S Q D L L K R Y Y R
ACAGTCAAAAATTGTTCAGGAAAGCCAAGATCTGCTGAAAAGATACTATA 841
H D D P E I G A H H L G F L W A
GGCATGACGATCCTGAAATAGGAGCACATCATCTTGGCTTTCTCTGGGCC 891
S L A N T I P A M F W A M Y Y I L
TCTCTAGCAAACACCATTCCAGCTATGTTCTGGGCAATGTATTATATTCT 941
R H P E A M E A L R D E I D S F L
TCGGCATCCTGAAGCTATGGAAGCCCTGCGTGACGAAATTGACAGTTTCC 991
Q S T G Q K K G P G I S V H F T
TGCAGTCAACAGGTCAAAAGAAAGGGCCTGGAATTTCAGTCCACTTCACC 1041
R E Q L D S L V C L E S T I L E V
AGAGAACAATTGGACAGCTTGGTCTGCCTGGAAAGCACTATTCTTGAGGT 1091
L R L C S Y S S I I R E V Q E D M
TCTGAGGCTGTGCTCATACTCCAGCATCATCCGAGAAGTGCAGGAGGATA 1141
N L S L E S K S F S L R K G D F
TGAATCTCAGCTTAGAGAGTAAGAGTTTCTCTCTGCGGAAAGGAGATTTT 1191
V A L F P P L I H N D P E I F D A
GTAGCCCTCTTTCCTCCACTCATACACAATGACCCGGAAATCTTCGATGC 1241
P K E F R F D R F I E D G K K K S
TCCAAAGGAATTTAGGTTCGATCGCTTCATAGAAGATGGTAAGAAGAAAA 1291
T F F K G G K R L K T Y V M P F
GCACGTTTTTCAAAGGAGGGAAGAGGCTGAAGACTTACGTTATGCCTTTT 1341
G L G T S K P G R Y F A V N E M
GGACTCGGAACAAGCAAATGTCCAGGGAGATATTTTGCAGTGAACGAAAT 1391
K L L L I E L L T Y F D L E I I D
GAAGCTACTGCTGATTGAGCTTTTAACTTATTTTGATTTAGAAATTATCG 1441
R K P I G L N H S R M F L G I Q
ACAGGAAGCCTATAGGGCTAAATCACAGTCGGATGTTTTTAGGTATTCAG 1491
H P D S A V S F R Y K A K S W R S
CACCCCGATTCTGCCGTCTCCTTTAGGTACAAAGCAAAATCTTGGAGAAG 1541
CTGAAAGTGTGGCAGAGAAGCTTTGCAGAGTAAGGCTGCATGTGCTGAGC 1591
TCCGTGATTTGGTGCACTCCCCCAAATGCAACCGCTACTCTTGTTTGAAA 1641
ATGGCAAATTTATATTTGGTTGAGATCAATCCAGTTGGTTTTGGGTCACA 1691
AAACCTGTCATAAAATAAAGCAGTGTGATGGTTTAAAAAATGTCATGGCA 1741
ATCATTTCAGGATAAGGTAAAATAACATTTTCAAGTTTGTACTTACTATG 1791
ATTTTTATCATTTGTAGTGAATGTGCTTTT 1821

must be concluded that recognition of the AUG codon by 40S ribosomal subunits is weak, resulting in weak initiation of Hct1 translation.

As predicted, sequence homology to the mouse cDNA for Hct1 is highest for human cholesterol 7α- hydroxylase, with 55.9% identity over 1054 bp of nucleic acids (homologies to be presented in more detail in Chapter Five). At the homology criteria employed, database homology searches failed to reveal homology to other cytochromes P-450. Thus, Hct1 may represent the second member of the *CYP7* family of cytochromes P-450.

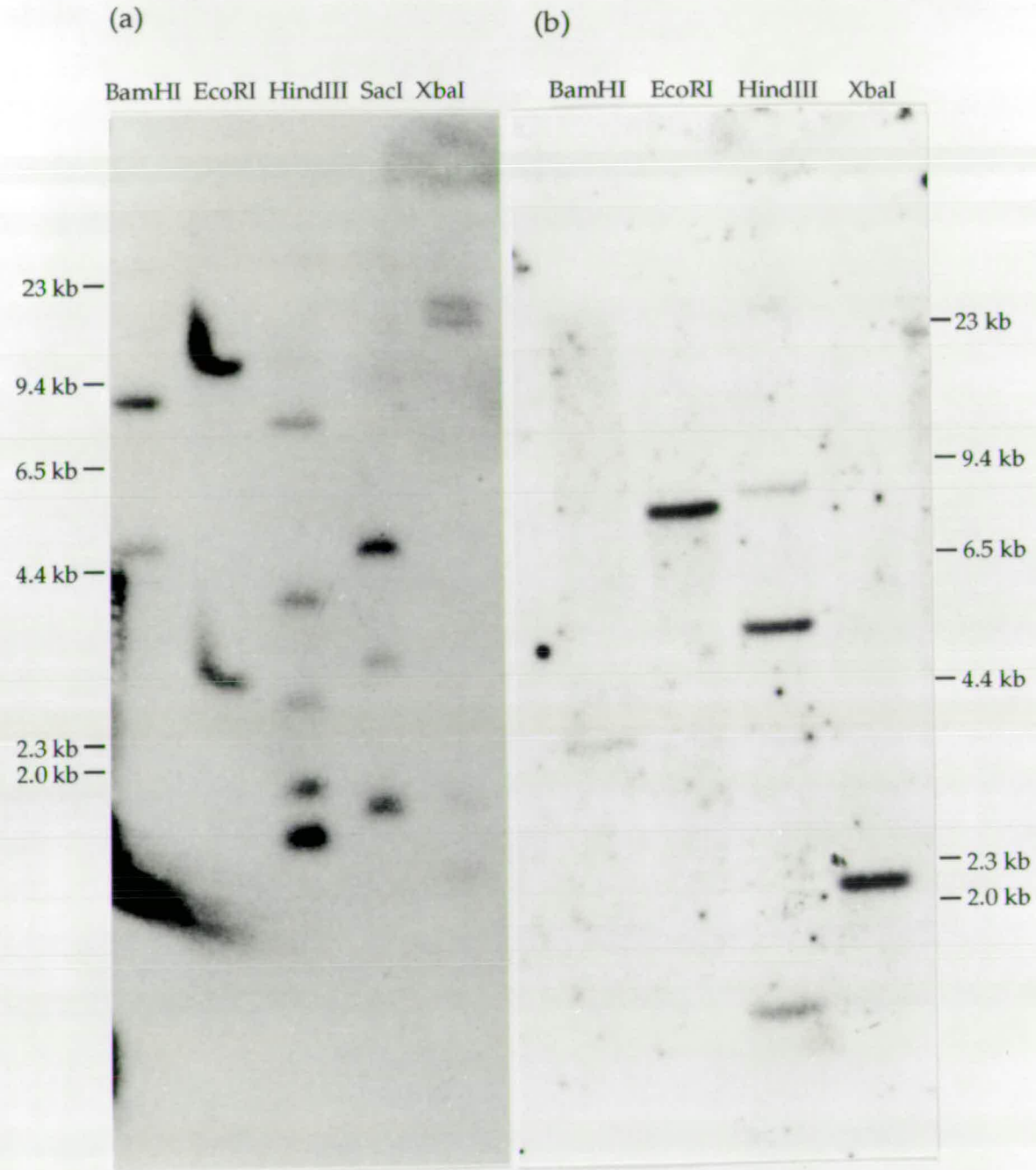
4.4 The structure of the mHct1 gene.

The use of homologous recombination to manipulate the mouse Hct1 gene requires knowledge of the intron-exon structure of the gene. Sequences upstream of the first Hct1 exon could also be analysed for elements which contribute to the transcriptional regulation of Hct1 expression. For these reasons, the organisation of the mouse Hct1 gene was investigated.

To assess the complexity of the Hct1 gene in the genome, that is, whether the Hct1 gene is present as a single copy in the haploid mouse genome, and to assist in mapping of mHct1 phage clones, the 1.8 kb full length mouse Hct1 clone was ^{32}P-labelled by random primer labelling and used as a probe on a Southern blot of mouse genomic DNA (Figure 4.4(a)). Under high stringency conditions the Hct1 probe recognised a small number of bands within the mouse genomic digests, suggesting that Hct1 is present in the mouse genome as a single copy gene. To confirm this, the original 0.3 kb cDNA clone, 14.5a, was used to probe a rat genomic Southern blot. The smaller probe hybridised to a single band in BamHI-, EcoRI-, and XbaI -digested genomic rat DNA (Figure 4.4(b)).

A mouse genomic DNA library (a gift from A. Reaume, Toronto) prepared from ES cells derived from the 129 mouse strain was screened for genomic clones containing mHct1 exonic sequence. 750,000 recombinant phage of the *lambda* DASH II library were plated at a

Figure 4.4 Genomic DNA Southern blot analysis of Hct1. (a) Mouse genomic DNA probed with the full-length mouse Hct1 cDNA clone. (b) Rat genomic DNA probed with clone 14.5a (original 0.3 kb clone of rHct1). 10 μg of genomic DNA was digested with the indicated enzymes.


(a)
BamHI EcoRI HindIII SacI XbaI
23 kb
9.4 kb
6.5 kb
4.4 kb
2.3 kb
2.0 kb
(b)
BamHI EcoRI HindIII XbaI
23 kb
9.4 kb
6.5 kb
4.4 kb
2.3 kb
2.0 kb

density of 50,000 recombinants per 15 cm plate. Duplicate lifts were made and probed with the 1.4 kb rat Hct1 clone (the mouse clone was as yet unavailable). After the primary screen, 5 clones were isolated. After secondary screening, three of these phage clones were positive and were purified.

Small scale phage DNA was prepared from each phage lysate and cut with NotI to release the inserts. No internal NotI sites were found in any of the clones. Clone λ-2 contained a 14 kb insert; clone λ-6 contained a 15 kb insert, and clone λ-11 contained a 12 kb insert.

These phage clones were mapped by a combination of restriction enzymes which either cut the *lambda* clones rarely, or by using restriction sites found in the mHct1 cDNA sequence (Figure 4.3(a)). A 5' probe was created using a 200 bp fragment from the 5' end of mHct1 cDNA as a probe; this segment extended from the internal BamHI site to an EcoRI site located in the polylinker. The 200 bp 3' cDNA probe extended from the SacI site to the polylinker NotI site. Exon-intron boundaries were determined by subcloning of exon-containing genomic DNA fragments and sequencing (Figure 4.5).

Phage clones λ-6 and λ-11 represented 20 kb of contiguous sequence of the Hct1 locus. λ-2 does not overlap with λ-6 or λ-11, thus the map of the Hct1 gene in mouse is incomplete. However, the present map shows that mHct1 spans at least 25 kb of the genome. At least two exons are contained within λ-6. The first exon (referred to as exon II) contains 133 bp of coding sequence, followed by exon III, located 4.0 kb downstream. The 3' boundary of this latter exon is not defined, however approximately 400 bp downstream of its 3' boundary commences exon IV, which together comprise 797 bp of coding sequence. Exon III and IV are also represented in the overlapping sequence of λ-11. A fourth exon of at least 345 bp was identified in λ-2 (referred to as exon VI). The 3' boundary of this exon has not been identified, thus it is not known whether this contains the remaining coding sequence or if there are additional exons.

As Table 4.1 shows, cDNA sequence from nucleotides 1073 - 1246 is not represented in the identified exons and must be represented in a separate exon. 142 bp of 5' sequence and 227 bp of 3' sequence have not

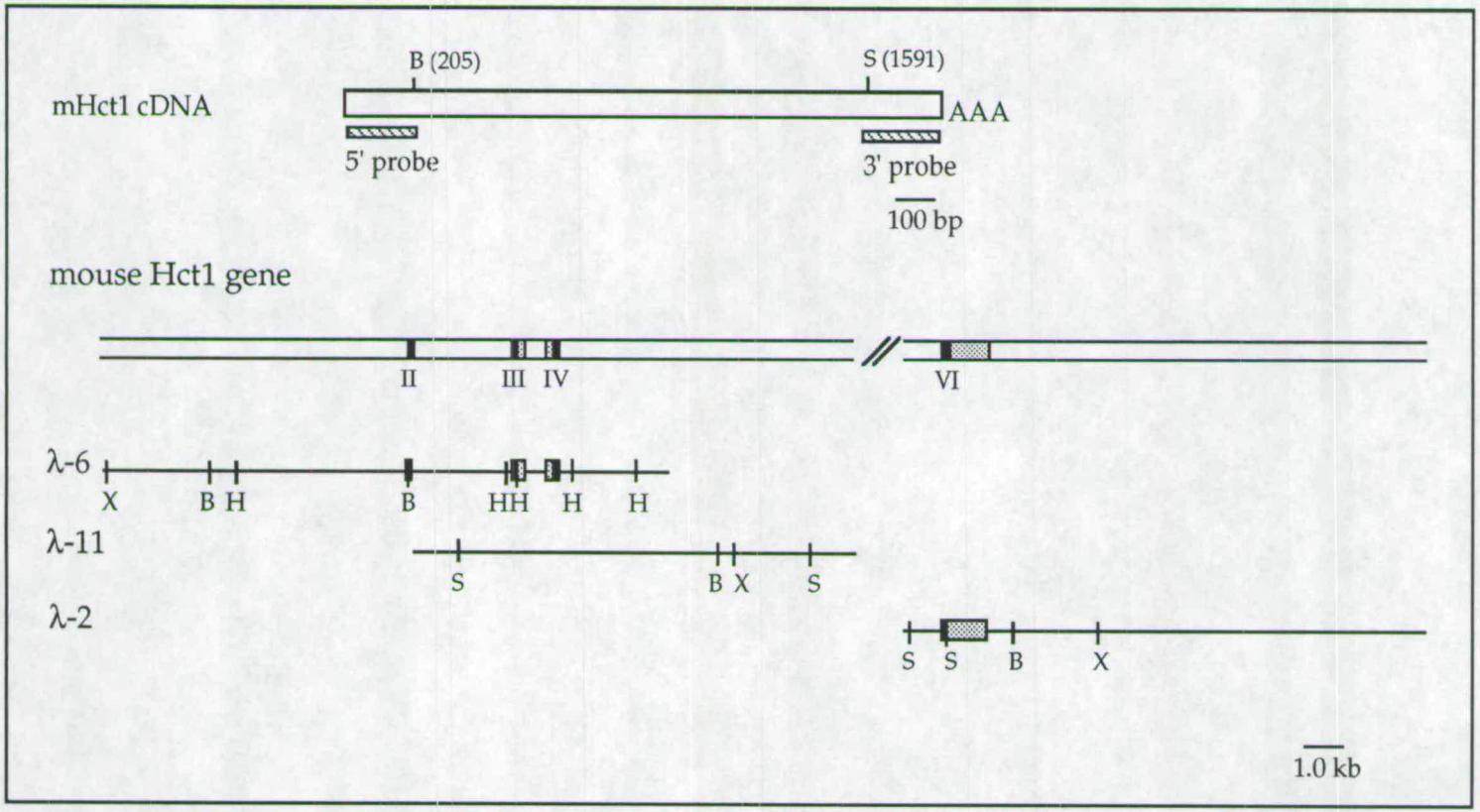


Figure 4.5 Genomic map of mouse Hct1 (incomplete). Exons II, III, IV and VI are represented on the phage clones (filled boxes). Exons I and V are not located. As indicated in Table 4.1, the boundaries of exons II, III and VI are not fully defined (stippled boxes).
B (BamHI); H (HindIII); S (SacI); X (XhoI)

yet been located in the genomic clones. The remaining 5' sequence is most likely contained in one exon, as the 5' probe (BamHI fragment) consistently recognised two bands by Southern analysis (one of which is exon II sequence) (data not shown). The remaining 3' sequence has not been located and may be part of exon VI or be encoded by a separate exon.

Exon	cDNA sequence represented	exon size (bp)	*CYP7* exon (bp)
I*	1 - 142	142	144
II	143 - 275	133	241
III	276 - ?	⎫ 797	587
IV	? - 1072	⎭	131
V*	1073 - 1246	174	176
VI**	1247 - (1821)	(575)	1596

Table 4.1 A summary of the exon-intron structure of Hct1 (incomplete) and comparison to human *CYP7* gene structure. * indicates that these exons are not cloned and are not necessarily one exon. ** indicates that the 3' boundary of exon VI is not confirmed and may not necessarily be the final exon. Human *CYP7* exon information is from Nishimoto *et al.* (1993).

4.5 The 5.0 kb transcript associated with rat Hct1

The 3' sequence of clone 13 was clearly distinct from the 3' ends of either clone 5 or clone 12. The original 300 base pair 14.5a clone that recognised the 5.0 kb transcript was used to isolate clone 13 (2.5 kb), indicating that clone 13 also derives from the Hct1 transcription unit. Because clone 13 had a unique 3' end, this link was also assumed to be through the use of an alternative polyadenylation site or through alternative splicing of a 3' end. To confirm the correspondence between

clone 13 and the 5.0 kb mRNA, clone 13 was used as a probe for Northern analysis (Figure 4.6(a)). A single 5.0 kb transcript only was recognised in all brain subregions tested. Clone 13 never detected the 1.8 kb or 2.2 kb transcripts even after prolonged autoradiographic exposure.

A 42-mer oligonucleotide designed from clone 13 sequence, situated 112 bases upstream from the start of the polyA tail, was used to examine clone 13 expression in sections of the adult rat brain by *in situ* hybridisation (Figure 4.6). The pattern of expression of clone 13 did not resemble that of clone 12 (Hct1) (refer to chapter 3, Figure 3.2). mRNA levels detected by the clone 13-specific oligonucleotide were very high compared to Hct1 clone 12 probe. Hct1 (clone 12) expression in the brain required exposure times of 6-8 weeks, while clone 13 gave a strong signal in 2-3 weeks. Expression of clone 13 appeared highest in the hippocampus and the habenulae (Figure 4.6(b)). Within the hippocampus, expression was highest in the dentate gyrus with uniform expression of clone 13 through the CA fields (Figure 4.6(e). Elsewhere in the adult brain clone 13 expression was widespread and uniform (Figure 4.6(c) (d)), compared to the very restricted pattern of expression observed for Hct1 clone 12 (Figures 3.2 and 3.3).

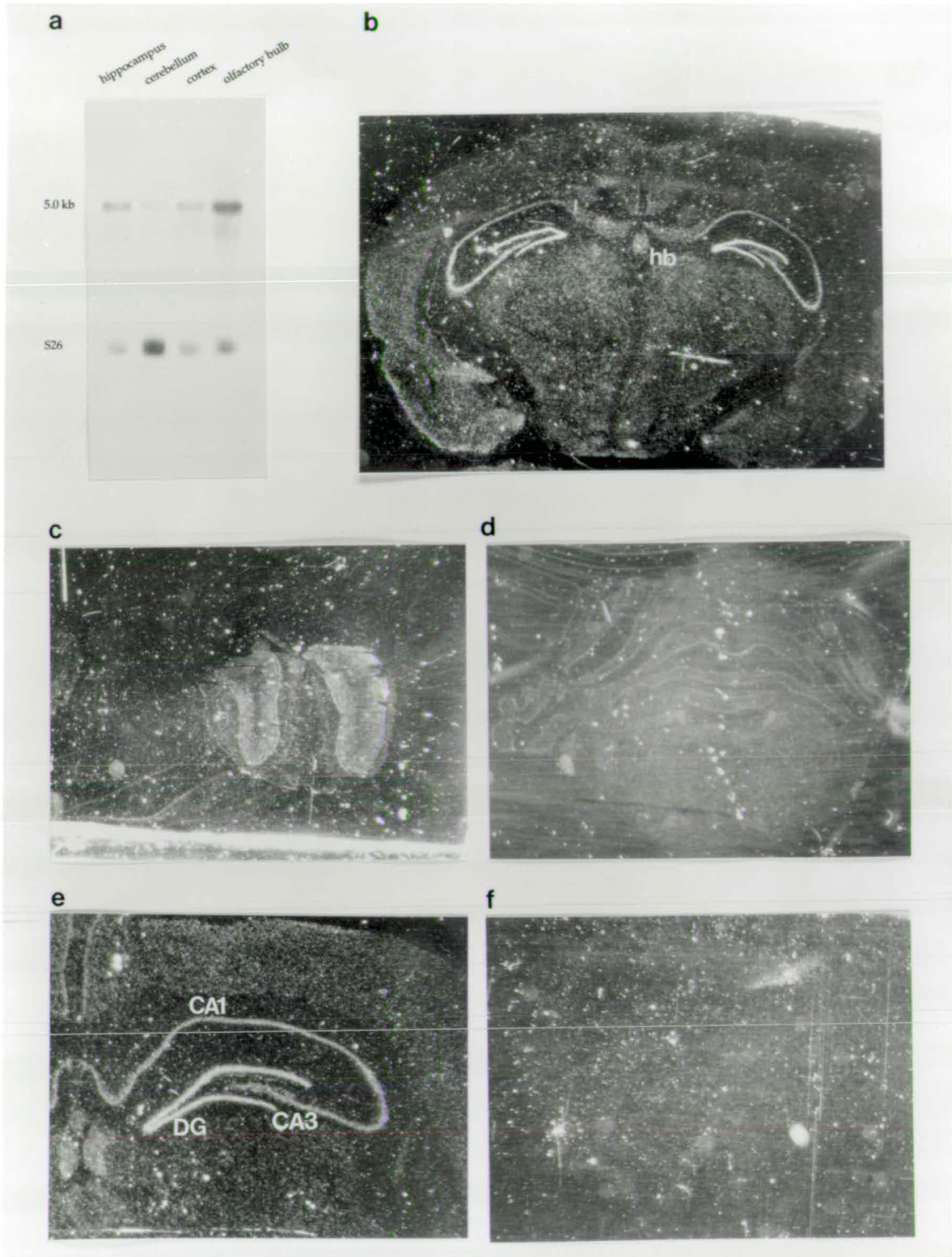
The fact that clones 12 (Hct1) and 13 were initially isolated using the same probe suggested that they were highly related, however restriction mapping failed to reveal any relationship. To explore this further, the hippocampus *lambda* ZAP II cDNA library was rescreened using clone 13 as a probe. Eight clones were purified after the secondary screen, and on plasmid rescue into pBluescript II SK, four were shown to contain inserts. Sequencing the ends of each showed that the four clones shared a common 3' end, though two had a 20 bp insert between the common 3' sequence and the polyA tail (Figure 4.7). Because the RNA used to make the hippocampus cDNA library was prepared from the hippocampus of several rats, this difference in 3' ends could result from allelic variation.

Figure 4.6 Expression of clone 13 (Hct2) in adult rat.

(a) Northern analysis of clone 13 expression in subregions of adult rat brain.

(b) *In situ* hybridisation of an Hct2-specific radiolabelled oligonucleotide to 10μ coronal sections of adult rat brain. Dark field view of Hct2 signal in rat brain coronal section through the hippocampus (magnification x7.5).
(c) Dark field view of Hct2 hybridisation signal in a coronal section of olfactory bulb (magnification x7.5).
(d) Dark field view of Hct2 signal in a coronal section of cerebellum (magnification x7.5).
(e) Distribution of Hct2 hybridisation signal in the hippocampus (magnification x20).
(f) Dark field view of rat coronal section through the hippocampus hybridised with an oligonucleotide specific for opsin (negative control).

CA1, CA3 pyramidal cell layers; DG, dentate gyrus; hb, habenulae


a
hippocampus
cerebellum
cortex
olfactory bulb
5.0 kb
S26
b
hb
c
d
e
CA1
DG
CA3
f

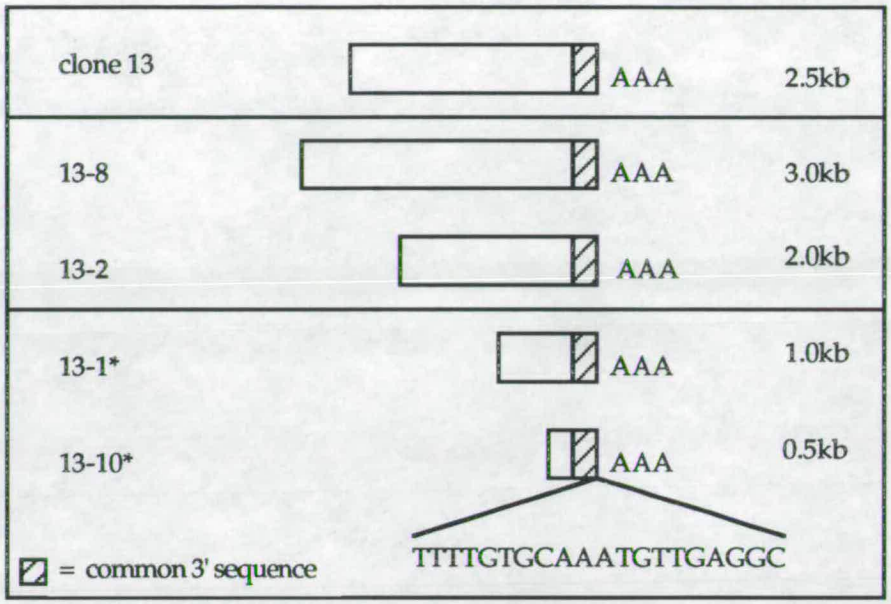


Figure 4.7 Subclones of clone 13. Insert sizes were determined by EcoRI, NotI digests of the recombinant plasmids. The 3' sequence is common to all clones except for clones 1 and 10 (*) which had an extra 20 bases inserted before the polyA tail.

The longest of the subclones, 13-8, (3.0 kb) was selected for sequencing. Exonuclease III deletions were made in the cDNA, the subclones sequenced directly using a primer hybridising to the T7 polymerase site, and the resulting contig sequences assembled (Figure 4.8).

The partial nucleotide sequences of clone 13-8 contigs were compared to known sequences contained in the GenBank/EMBL database. The recently submitted sequence of the human ABR cDNA (Tan *et al*, 1993) was found to be very similar to clone 13. ABR (active BCR-related) is a gene encoding a GTPase-activating protein (GAP), which acts on specific members of the Ras superfamily of small GTP-binding proteins. However, while regions of clone 13 were virtually identical to ABR, other stretches of clone 13 were completely unrelated to ABR (Figure 4.8). As 13-8 appears to be more closely related to ABR than to Hct1, this clone, and its corresponding gene, was termed Hct2.

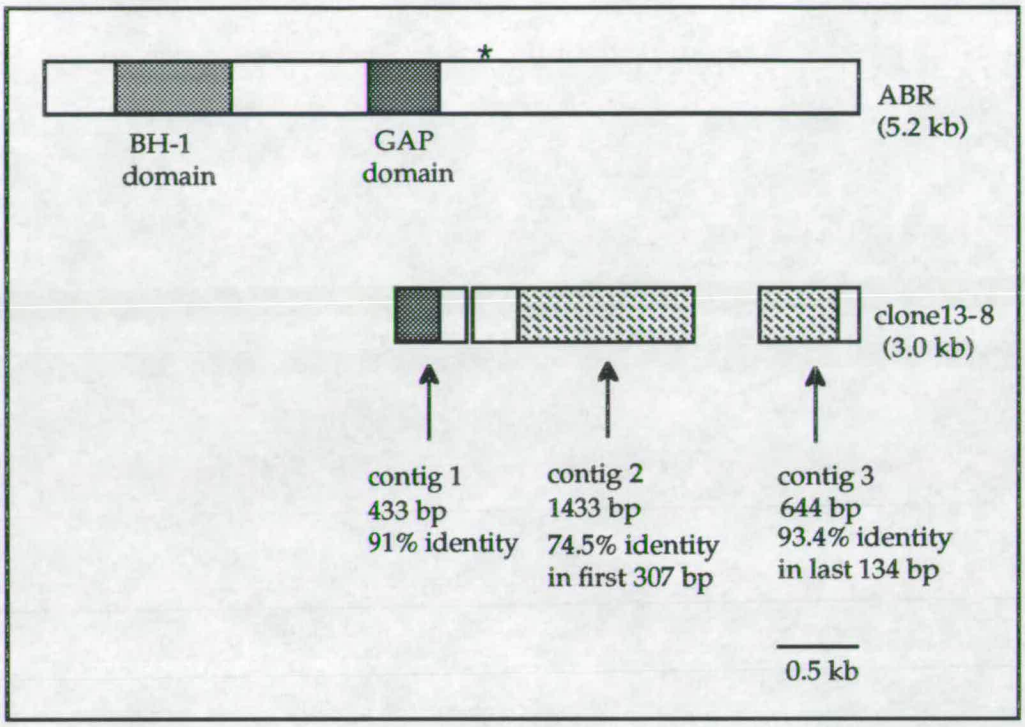


Figure 4.8(a) Regions of conservation between Hct2 (clone 13-8) and human ABR.
The white boxed areas in clone 13-8 represent regions in common between the two genes. Hatched regions represent non-homologous regions between clone 13-8 and ABR. All of contig 1 was 91% identical to ABR. The first 307 bp of contig 2 was 74.5% identical, and the last 134 bp of contig 3 was 93.4% identical to human ABR. The stippled boxes represent mapped functional domains in ABR, part of which is also found in clone 13-8. The asterisk indicates the position of the stop codon(TAG) at nucleotide 2688.

(b) Amino acid alignment of the racGAP activity domain of human ABR with part of the same domain identified and conserved in clone 13-8.

```
Hct2
ABR  (621)  CVEEVEKRGIEEVGIYRISGVATDIQALKAVFDANNKDILLMLSDMDINAI
consensus   C  EIE RG    G YRVSG       L   FD            D N


Hct2              YFRELPEPLLTDRLYPAFMEGIALSDPAAKEICMMMHLLRSLPDP
ABR         AGTPKLYFRELPEPLLTDRLYPAFMEGIALSDPAAKENCMM-HLVRSLPDP
consensus   AG  K YFRELP PL T   Y  F E     DP A            LP


Hct2        NLITFLFLLEHLKRVAEKEPINKMSLHNLATVFGPTLLRPSE
ABR         NLITFLFLLEHLKRVAEKEPINKMSLHNLATVFGPTLLRPSE
consensus   N  T  YL  HL RVAE   VN M   NLA VFGPTL R  E
```

This unusual pattern of homology between human ABR and rat Hct2 (clone 13-8) might be explained in at least two ways. It is possible that the non-homologous region is a consequence of the cDNA cloning and that clone 13-8 is an artefact. Alternatively, this could be a previously undescribed form of ABR which has been generated through an alternative splicing event. In an attempt to resolve the cDNA structure of Hct2 (clone 13), the remaining subclones, numbered 1, 2 and 10, were sequenced from their 5' ends. These 5' sequences of clones 1 and 10 were compared with the GenBank/EMBL database sequences, and were shown to be highly homologous to the human ABR gene (Figure 4.9). The sequenced ends of clone 2 aligned within the ABR sequence of clone 13.

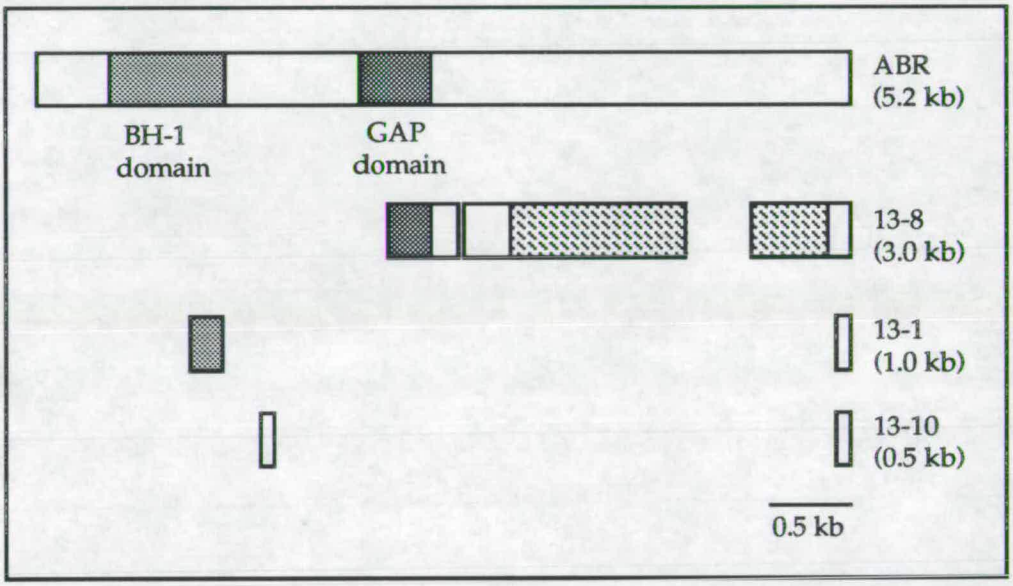


Figure 4.9 Comparison of human ABR cDNA clone with partial sequences from clone 13 subclones 8 (3.0 kb), 1 (1.0 kb), and 10 (0.5 kb). Sequence within the subclones which is homologous to ABR is indicated. Non-homologous regions within subclone 8 corresponds to the hatched boxes. Stippled boxes in the ABR cDNA locate the 5' BH-1 or dbl-like domain and the 3' GAP domain.

Clones 1 and 10, which are 1.0 kb and 0.5 kb respectively, aligned within the ABR gene in regions which do not correlate with their size.

Based on size, both clones would be expected to align within the 3' untranslated region. However, the 0.5 kb clone 10 aligned 5' of the conserved GAP domain, while clone 1 aligned within the BH-1 domain (Figure 4.9). These data suggest that the cDNA clones obtained from the hippocampus cDNA library are most likely generated through alternative splicing. The possibility that clone 8 is artefactual has not been ruled out, however two independent clones (1 and 10) support differential transcripts.

4.6 Discussion

Transcription of the rat Hct1 gene generates two transcripts through the use of alternative polyadenylation sites. These two transcripts (1.8 kb and 2.2 kb) were present in all tissues examined. A third transcript can also be seen by Northern analysis, however this is not related to the Hct1 transcription unit. Clone 13 encodes a novel message of the rat ABR gene, where the 5' and 3' ends encode the rat ABR transcript, but inserted into this is a novel sequence in place of the ABR 3' untranslated region. Smaller subclones of clone 13 contained ABR sequence but not at the predicted position relative to their size. These isolated cDNA clones of rat ABR predict that alternative splicing of the mRNA produces different messages. Although these alternative forms of ABR mRNA are not detected by Northern analysis, it may be that these messages are less abundant and therefore not detectable by this method. Confirmation of these conclusions would require further analysis of the subclones and of genomic clones of the rat ABR gene. Rat ABR (clone 13) was isolated in a screen using rHct1 (clone 14.5a) as a probe. Exactly how ABR was isolated from such a screen is not clear. No significant homologies between the 0.3 kb probe and the partially sequenced clone 13 can be found.

A full length cDNA for mouse Hct1 would be predicted to encode the 1.8 kb transcript, consistent with the size of most cloned cytochromes P-450, where the transcript size varies between 1.8 - 2.2 kb (Paine, 1991). No Hct1 cDNA clones longer than 1.8 kb, which would be expected to encode the 2.2 kb transcript, were isolated from the mouse

liver cDNA library. This is supported by mouse Hct1 Northern analysis which only detects the 1.8 kb transcript (Figure 3.5).

The consequences of the use of two alternative polyadenylation sites for Hct1 are not known, however the 3' untranslated regions of certain mRNAs have been shown to influence both the stability of the transcript and the translation of the RNA (for review, see Sonenberg, 1994). The human amyloid protein precursor (APP) gene, for example, produces two transcripts which are generated through the use of two alternative polyadenylation sites. de Sauvage *et al.* (1992) have shown that sequences contained between the two polyadenylation sites cause a 3-fold increase in translation of APP mRNA.

Analysis of the genomic structure of Hct1 suggests that the gene contains at least six exons, although analysis is as yet incomplete. Because most cytochrome P-450 gene families have relatively conserved gene structures (Paine, 1991), *CYP7* and Hct1 should have a similar gene structure if they are members of the same gene family. The human *CYP7* gene spans 17 kb, comprising 6.0 kb of 5' flanking sequence and 10.5 kb containing the entire transcription unit which consists of six exons (Nishimoto *et al.*, 1993). Most cytochrome P-450 genes are encoded in 7 - 13 exons, however *CYP7*, and possibly Hct1 which has a minimum of six exons, is encoded by a smaller number of exons. Interestingly, most CYP genes including *CYP7*, encode the haem-binding domain(a conserved C-terminal domain (Chapter 5) and the 3' untranslated region together in the last exon. Based on this, it might be predicted that exon VI is the final exon, although this remains to be confirmed. The isolation of further genomic clones to complete the Hct1 map will help to clarify the relationship between Hct1, *CYP7* and other cytochrome P-450 genes. However, although Hct1 spans a much larger piece of the genome, as Table 4.1 shows, the emerging pattern of Hct1 exons is similar to the *CYP7* gene.

A cDNA for rat ABR (active bcr-related) was isolated. Rat ABR expression is widespread in the adult brain and thus is not relevant to the main objective of this study, but the potential function of ABR in the brain is of considerable interest. ABR is a GTPase activating protein (GAP) with activity specific for rac, which is a member of the rho

family of small GTP-binding proteins (of which ras is the most well known) (reviewed in Downard, 1990). Rac and Rho belong to a subfamily of Ras-related proteins and their function is one of cytoplasmic signalling, acting to couple extracellular signals with the cytoskeleton (Hall, 1993). Thus, for example, activation of rac in fibroblasts leads to changes in cell morphology by activating actin polymerisation (Ridley *et al.*, 1992). GAPs act as regulatory molecules on the activity of Ras-related proteins by converting them to an inactive state, and the GAP activity of ABR has been shown specifically to convert rac to an inactive state.

The ABR cDNA clones described, while requiring confirmation, could be generated by the proposed model of alternative splicing described in Figure 4.10. Interestingly, if these cDNAs are not cloning artefacts, such splicing products would produce forms of ABR which lack the GAP domain of the protein. The role of such proteins in rac signalling would prompt further investigation. One possibility is that these proteins could bind to rac, but not facilitate the conversion of GTP to GDP, maintaining rac in an active, GTP-bound state. Alternatively, the BH-1 domain, which has homology (but as yet no demonstrated activity) to guanine nucleotide exchange factors, is maintained in these cDNA clones. The function of these "GAPless" cDNA clones could be determined by the activity of the BH-1 domain.

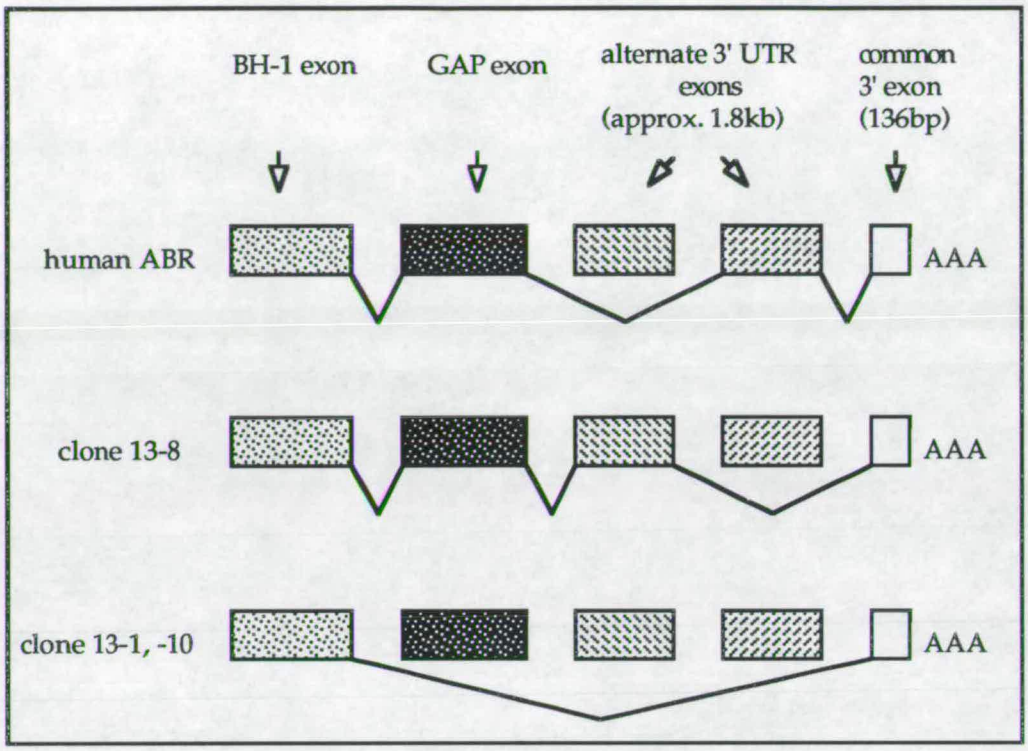


Figure 4.10 A model of alternative splicing proposed to explain the observed cDNA subclones of rat ABR.
The common 3' sequence for all clones is 136 bp. The novel 3' UTR sequence is approximately 1.8 kb. The GAP domain itself is 447 bp. Subclones 1 and 10 are 1.0 kb and 0.5 kb respectively, and subclone 13-8 is 3.0 kb.

The role of Rac and ABR in the brain is unknown, however the neuronal cytoskeleton has been suggested to participate in synaptic plasticity (Lynch and Baudry, 1987). In addition, the effects of cytoskeletal elements have been shown on the localisation of nicotinic acetylcholine receptors at the neuromuscular junction (Froehner *et al.*, 1991). More recently, the state of actin polymerisation has been shown to affect the activity of the NMDA receptor channel (Rosenmund and Westbrook, 1993). It may be of future interest to look at a possible role for Rac signalling in effecting postsynaptic changes of the neuronal cytoskeleton, in particular, at the postsynapse.

CHAPTER FIVE

Hct1 STRUCTURE AND FUNCTION

Introduction

The analysis of cDNA and deduced amino acid sequences demonstrate that Hct1 is a previously undescribed member of the cytochrome P-450 superfamily. As an initial step towards characterising a possible role for Hct1, evidence for conservation of the gene in vertebrates, and for conserved domains within Hct1 and other cytochromes P-450 was investigated.

Cytochromes P-450 are structurally conserved, and are divided into thirty-six distinct families based on sequence homology. While the range of functions carried out by cytochromes P-450 is diverse, they all carry out essentially the same enzymatic reaction, and this is reflected in conservation of protein structure. Structurally conserved domains in a protein generally identify regions important for the function of the protein, while slight variations within such a domain can provide further classification of related proteins. The identification of the homeodomain in *Drosophila* genes, for example, led to the identification of a large family of transcription factors conserved throughout evolution (Gehring, 1987). Cytochromes P-450 are evolutionarily conserved, with representatives from bacteria to human. However, not all individual cytochromes P-450 are conserved between vertebrate species. For example, the rat *CYP2D* subfamily has five active genes, while the same subfamily in human has only one active gene and two pseudogenes (Nelson *et al*, 1993).

The enzymatic activity of Hct1 protein was also addressed. In order to assay for cytochrome P-450 activity on a proposed substrate, the protein must first be expressed in a suitable system where the

coenzyme, cytochrome P-450 reductase, is also expressed. A yeast expression system was chosen both for ease of manipulation and because yeast express an endogenous form of cytochrome P-450 reductase. The successful use of yeast as an expression system for cytochromes P-450 has been described (for examples see Black *et al*, 1989; Ching *et al*, 1991).

5.1 Conservation of Hct1 in humans.

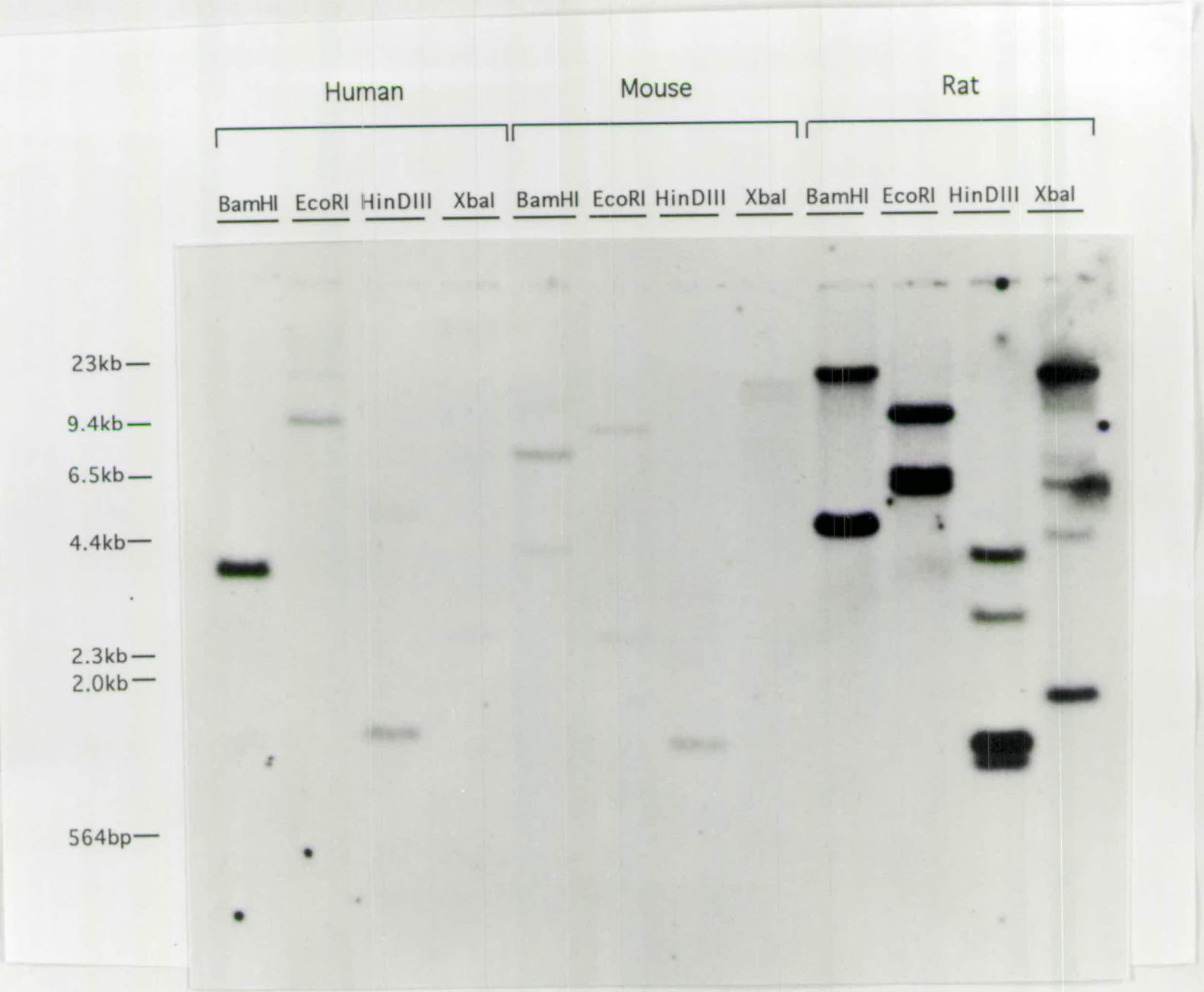
The evolutionary conservation of a gene supports a functionally significant role for that gene in the organism. The conservation of Hct1 in rodents has been demonstrated by the cloning of the rat and mouse cDNAs for Hct1. To establish the presence of the Hct1 gene in the human genome, Southern blotting of human DNA was performed. The rat 1.4 kb clone of Hct1 was used as a radiolabelled probe and gave strong signals from all three species (Figure 5.1). A number of hybridising fragments appear to be conserved between species, suggesting conservation of the Hct1 gene structure. There is a conserved 1.4 kb HindIII band conserved between mouse and rat, while human DNA contains a slightly larger HindIII band of 1.6 kb. Also an EcoRI fragment of 11 kb is conserved in human and rat Hct1. Conservation of Hct1 gene structure is also supported from the cDNA digestion patterns of mouse and rat (see Figures 4.4(a) and 4.2(a)), where the SacI, HindIII and PstI sites are conserved between the rodent species.

5.2 Structure of mouse Hct1 protein

The mouse Hct1 open reading frame encodes a protein of 507 amino acids. Mouse and rat translated sequences show 81% similarity over 414 amino acids using the Lipman-Pearson algorithm (DNAstar), but the absolute value may be slightly different as the full protein sequence for rat Hct1 is not known. Nevertheless, this order of similarity between the mouse and rat is consistent for orthologous genes. Amino acid identity between mouse and rat cholesterol 7α-hydroxylase is of a similar order of similarity (Figure 5.3).

Figure 5.1 Genomic Southern of mouse, rat and human DNA probed with the rat 1.4kb clone for Hct1.

10 μg of DNA from each species was digested with BamHI, EcoRI, HindIII and XbaI; restriction enzymes known to cut rarely in the genome, and which are not affected by DNA methylation.


Human
Mouse
Rat
BamHI
EcoRI
HinDIII
XbaI
BamHI
EcoRI
HinDIII
XbaI
BamHI
EcoRI
HinDIII
XbaI
23kb
9.4kb
6.5kb
4.4kb
2.3kb
2.0kb
564bp

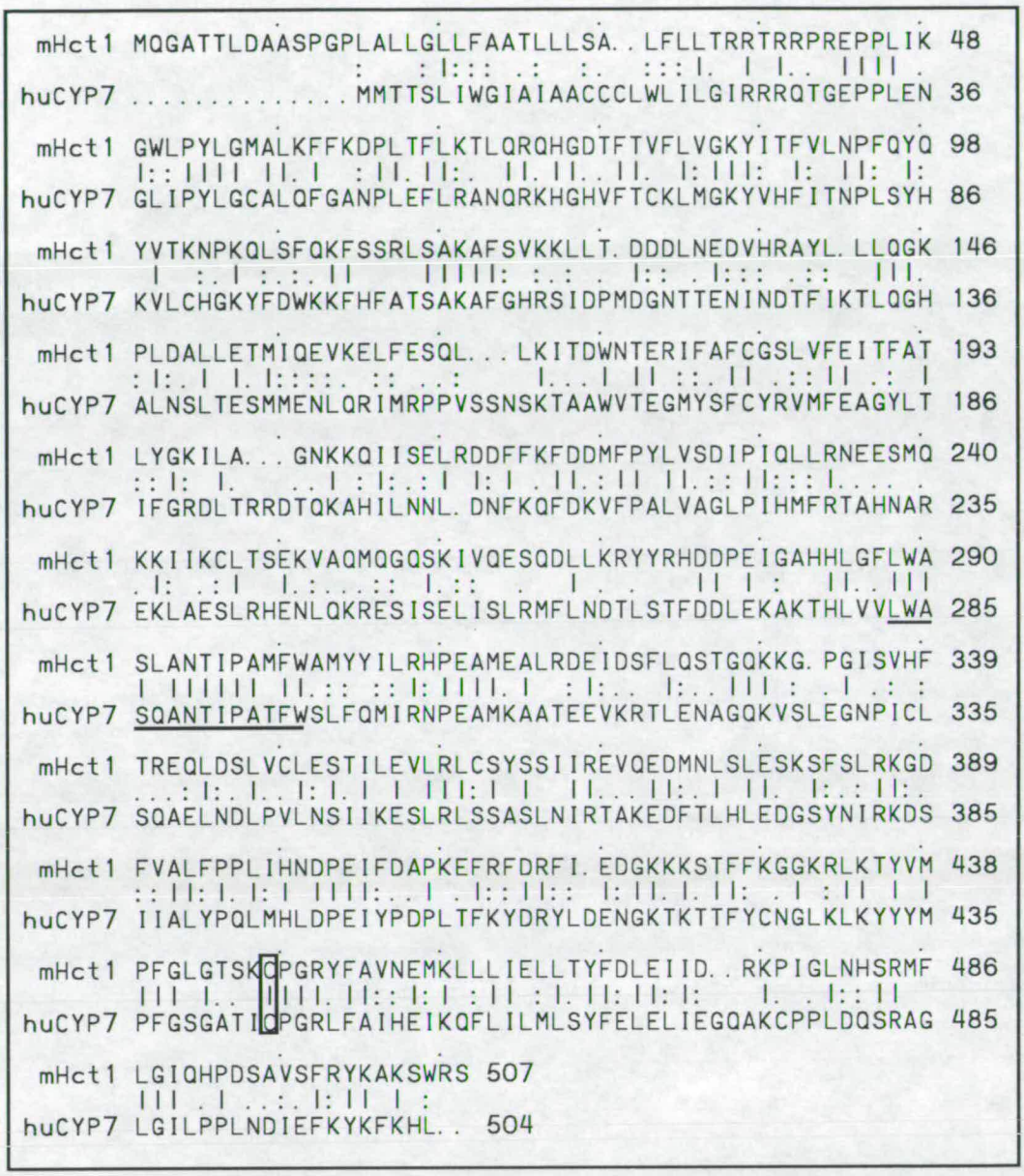


Figure 5.2 Amino acid alignment of mouse Hct1 and human CYP7. Identical amino acids are indicated by a bar, Conserved amino acid changes are indicated by (:). Dots in the sequence indicate introduced gaps in the alignment. The conserved cysteine (amino acid 447 in mHct1) in the haem-binding domain is boxed. Another domain of strong identity (amino acids 288 - 301) is underlined. The alignment was constructed using the GAP programme (UWGCG) set at default values.

Evidence for post-translational modifications of mHct1, through conservation of specific motifs was considered. Other cytochromes P-450, such as cholesterol 7α-hydroxylase, are known to contain potential glycosylation sites (Li *et al*, 1991). Two N-linked glycosylation sites were found in mHct1 protein. The consensus motif for N-linked glycosylation is Asn-X-Thr/Ser (where X is any amino acid except proline) (Marshall, 1972). Asparagines at amino acids 375 and 481 are potential glycosylation sites on mHct1. N-linked glycosylation sites are also found in cholesterol 7α-hydroxylase, however the positions between cholesterol 7α-hydroxylase and mHct1 are not conserved (Li *et al*, 1990).

Both human and rat cholesterol 7α-hydroxylase are 39% identical to mHct1 at the amino acid level (Figures 5.2 and 5.3). The decision as to whether or not a new cytochrome P-450 constitutes a new gene family is based on sequence homology. This will ultimately be decided by the P-450 Nomenclature Committee, however, in general, two cytochrome P-450 proteins are classed as belonging to the same family if they exhibit greater than 40% identity (Nebert *et al*, 1989). It is not clear therefore whether Hct1 will represent a new family of cytochrome P-450 genes, or a new subfamily of *CYP7* genes.

	mHct1	rHct1*	hu CYP7	rat CYP7
mHct1	-			
rHct1*	81%	-		
huCYP7	39.3%	37%	-	
rat CYP7	38.9%	36.7%	82%	-

Figure 5.3 Amino acid identity between Hct1 and the nearest homologue, cholesterol 7α-hydroxylase. (*) indicates that the rat Hct1 clone was not full length. Comparisons were made using the GAP programme (UWGCG).

Because the cytochrome P-450 superfamily is so large, a number of comprehensive alignment studies have highlighted conserved regions within cytochrome P-450 proteins (Nelson and Strobel, 1988). Thus, Hct1 was aligned with known regions of conservation between a cytochrome P-450 from *Psuedomonas putida*, CYP101 (the only P-450 for which crystallographic structural information is known). In addition, mammalian cytochromes P-450 which are known to act in steroidogenic pathways or modify fatty acids were aligned with mouse Hct1 protein (Figure 5.4).

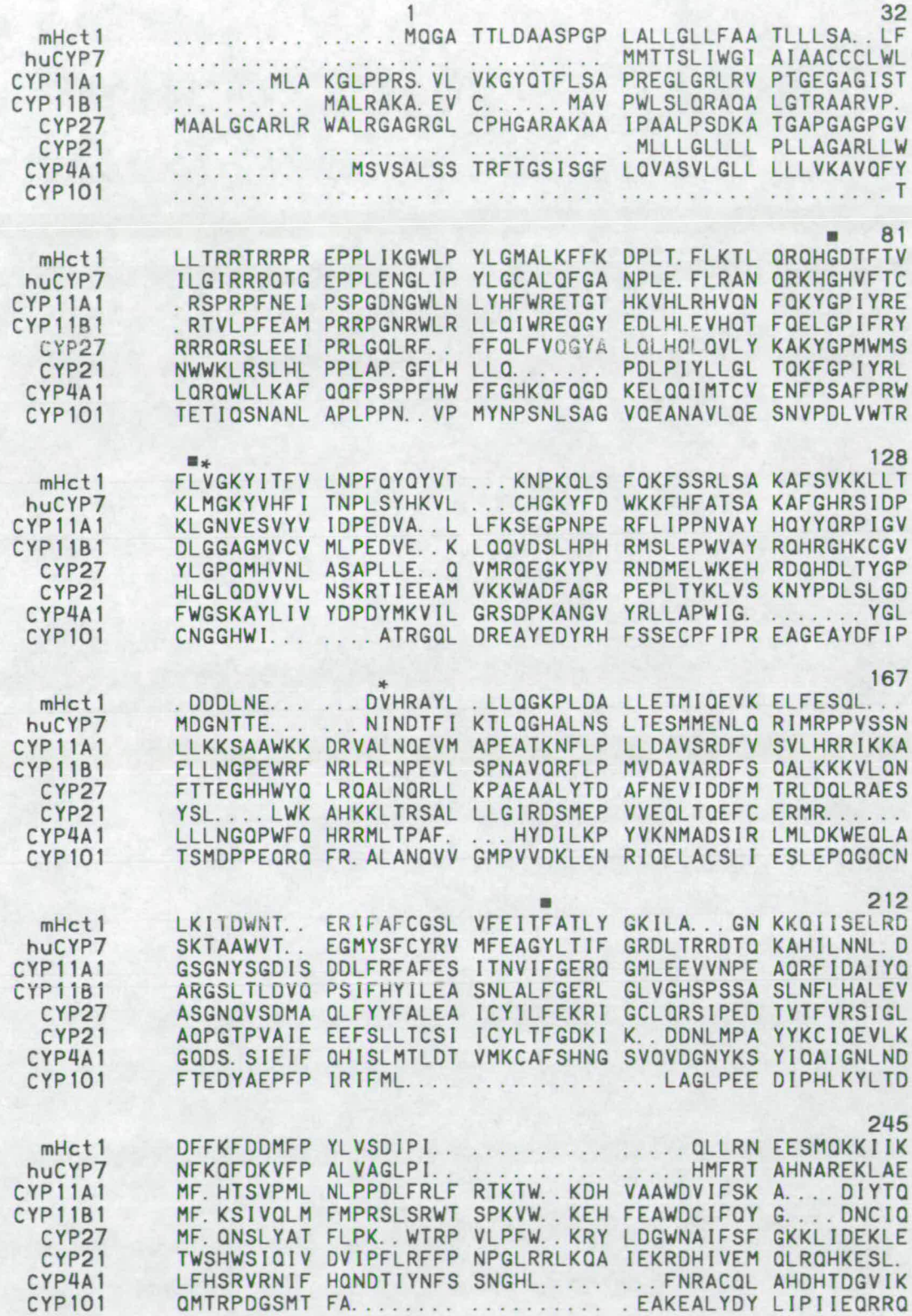
The alignment of cytochromes P-450 in Figure 5.4 identifies a number of regions of strong amino acid conservation between mHct1 and the selected CYP proteins. The cysteine at residue 447 is absolutely invariant in all cytochromes P-450. Other conserved amino acid positions are glu-356, leu-358, arg-359, pro-408, arg-415, phe-440, gly-443, gly-449, ala-453, and glu-456. There are also a number of conservative amino acid substitutions. Thus, the amino acids surrounding pro-400 and arg-407 tend to be aromatic (residues 397, 403 and 408), suggesting a conserved structure of cytochromes P-450 in this region.

Figure 5.2 demonstrates the similarity of CYP7 and mHct1 proteins. Both proteins are approximately the same length, around 500 amino acids, and the degree of identity is broadly equal over the entire length of the proteins. The amino acids from 288-301 are strikingly conserved, and as shown in Figure 5.4, do not appear to be conserved outside of these two proteins. There are regions where mHct1 and CYP7 are different from otherwise conserved regions of cytochromes P-450. In Figure 5.4, a number of amino acid deletions appear to be conserved between the two proteins. Amino acid deletions occur after amino acids 101, 134, and seventeen amino acids after position 230 are not conserved in mHct1 and CYP7. Similarly, insertions of amino acids

Figure 5.4 Alignment of cytochromes P-450 with mouse Hct1.
8/8 amino acids conserved in the CYP proteins presented (•);
7/8 amino acids conserved in the CYP proteins presented (∘);
amino acids conserved in steroidogenic CYP proteins (▪);
amino acid conserved except in mHct1 and CYP7 (*); and the
non- conserved glycine at position 283 (!).

The invariant cysteine residue at position 439 is highlighted in bold. Columns of amino acids (in bold) at positions 397, 403 and 408 indicate conserved aromatic amino acids. The highlighted amino acid motif (residues 280-293) indicates a conserved motif in mHct1 and CYP7 only.

The alignment was produced using the UWGCG programme PILEUP with all parameters set at default values.


1
32
mHct1 .... MQGA TTLDAASPGP LALLGLLFAA TLLLSA..LF
huCYP7 .... MMTTSLIWGI AIAACCCLWL
CYP11A1 ..... MLA KGLPPRS. VL VKGYQTFLSA PREGLGRLRV PTGEGAGIST
CYP11B1 ..... MALRAKA. EV C..... MAV PWLSLQRAQA LGTRAARVP.
CYP27 MAALGCARLR WALRGAGRGL CPHGARAKAA IPAALPSDKA TGAPGAGPGV
CYP21 .... .MLLLGLLLL PLLAGARLLW
CYP4A1 ....MSVSALSS TRFTGSISGF LQVASVLGLL LLLVKAVQFY
CYP101 ....T
81
mHct1 LLTRRTRRPR EPPLIKGWLP YLGMALKFFK DPLT.FLKTL QRQHGDTFTV
huCYP7 ILGIRRRQTG EPPLENGLIP YLGCALQFGA NPLE.FLRAN QRKHGHVFTC
CYP11A1 .RSPRPFNEI PSPGDNGWLN LYHFWRETGT HKVHLRHVQN FQKYGPIYRE
CYP11B1 .RTVLPFEAM PRRPGNRWLR LLQIWREQGY EDLHLEVHQT FQELGPIFRY
CYP27 RRRQRSLEEI PRLGQLRF.. FFQLFVQGYA LQLHQLQVLY KAKYGPMWMS
CYP21 NWWKLRSLHL PPLAP.GFLH LLQ..... PDLPIYLLGL TQKFGPIYRL
CYP4A1 LQRQWLLKAF QQFPSPPFHW FFGHKQFQGD KELQQIMTCV ENFPSAFPRW
CYP101 TETIQSNANL APLPPN..VP MYNPSNLSAG VQEANAVLQE SNVPDLVWTR
128
mHct1 FLVGKYITFV LNPFQYQYVT ...KNPKQLS FQKFSSRLSA KAFSVKKLLT
huCYP7 KLMGKYVHFI TNPLSYHKVL ...CHGKYFD WKKFHFATSA KAFGHRSIDP
CYP11A1 KLGNVESVYV IDPEDVA..L LFKSEGPNPE RFLIPPNVAY HQYYQRPIGV
CYP11B1 DLGGAGMVCV MLPEDVE..K LQQVDSLHPH RMSLEPWVAY RQHRGHKCGV
CYP27 YLGPQMHVNL ASAPLLE..Q VMRQEGKYPV RNDMELWKEH RDQHDLTYGP
CYP21 HLGLQDVVVL NSKRTIEEAM VKKWADFAGR PEPLTYKLVS KNYPDLSLGD
CYP4A1 FWGSKAYLIV YDPDYMKVIL GRSDPKANGV YRLLAPWIG. .....YGL
CYP101 CNGGHWI... ....ATRGQL DREAYEDYRH FSSECPFIPR EAGEAYDFIP
167
mHct1 .DDDLNE... ...DVHRAYL .LLQGKPLDA LLETMIQEVK ELFESQL...
huCYP7 MDGNTTE... ...NINDTFI KTLQGHALNS LTESMMENLQ RIMRPPVSSN
CYP11A1 LLKKSAAWKK DRVALNQEVM APEATKNFLP LLDAVSRDFV SVLHRRIKKA
CYP11B1 FLLNGPEWRF NRLRLNPEVL SPNAVQRFLP MVDAVARDFS QALKKKVLQN
CYP27 FTTEGHHWYQ LRQALNQRLL KPAEAALYTD AFNEVIDDFM TRLDQLRAES
CYP21 YSL....LWK AHKKLTRSAL LLGIRDSMEP VVEQLTQEFC ERMR.....
CYP4A1 LLLNGQPWFQ HRRMLTPAF. ...HYDILKP YVKNMADSIR LMLDKWEQLA
CYP101 TSMDPPEQRQ FR.ALANQVV GMPVVDKLEN RIQELACSLI ESLEPQGQCN
212
mHct1 LKITDWNT.. ERIFAFCGSL VFEITFATLY GKILA...GN KKQIISELRD
huCYP7 SKTAAWVT.. EGMYSFCYRV MFEAGYLTIF GRDLTRRDTQ KAHILNNL.D
CYP11A1 GSGNYSGDIS DDLFRFAFES ITNVIFGERQ GMLEEVVNPE AQRFIDAIYQ
CYP11B1 ARGSLTLDVQ PSIFHYILEA SNLALFGERL GLVGHSPSSA SLNFLHALEV
CYP27 ASGNQVSDMA QLFYYFALEA ICYILFEKRI GCLQRSIPED TVTFVRSIGL
CYP21 AQPGTPVAIE EEFSLLTCSI ICYLTFGDKI K..DDNLMPA YYKCIQEVLK
CYP4A1 GQDS.SIEIF QHISLMTLDT VMKCAFSHNG SVQVDGNYKS YIQAIGNLND
CYP101 FTEDYAEPFP IRIFML.... .....LAGLPEE DIPHLKYLTD
245
mHct1 DFFKFDDMFP YLVSDIPI.. .....QLLRN EESMQKKIIK
huCYP7 NFKQFDKVFP ALVAGLPI.. .....HMFRT AHNAREKLAE
CYP11A1 MF.HTSVPML NLPPDLFRLF RTKTW..KDH VAAWDVIFSK A....DIYTQ
CYP11B1 MF.KSIVQLM FMPRSLSRWT SPKVW..KEH FEAWDCIFQY G....DNCIQ
CYP27 MF.QNSLYAT FLPK..WTRP VLPFW..KRY LDGWNAIFSF GKKLIDEKLE
CYP21 TWSHWSIQIV DVIPFLRFFP NPGLRRLKQA IEKRDHIVEM QLRQHKESL.
CYP4A1 LFHSRVRNIF HQNDTIYNFS SNGHL..... ...FNRACQL AHDHTDGVIK
CYP101 QMTRPDGSMT FA..... .....EAKEALYDY LIPIIEQRRQ

```
                                                                    286
  mHct1     CLTSEKVAQM QGQSKIVQES QDLLKRY... ......YRHD DPEIGAHHLG
 huCYP7     SLRHENLQKR ESISELISLR MFLNDTL... ......STFD DLEKAKTHLV
CYP11A1     NFYWELRQKG SVHHDYRGML YRLL...... ....GDSKMS FEDIKANVTE
CYP11B1     KIYQEL..AF SRPQQYTSIV AELL...... ....LNAELS PDAIKANSME
  CYP27     DMEAQLQAAG PDGIQVSGYL HFLL...... ....ASGQLS PREAMGSLPE
  CYP21     .......... .VAGQWRDMM DYMLQGVAQP SMEEGSGQLL EGHVHMAAVD
 CYP4A1     LRKDQLQNAG ELEKVKKKRR LDFLDILLLA RMENGDS.LS DKDLRAEVDT
 CYP1O1     KPGTDAISIV ANGQV..... .......... ....NGRPIT SDEAKRMCGL

            !  *o                 o          o                    335
  mHct1     FLWASLANTI PAMFWAMYYI LRHPEAMEAL RDEIDSFLQS TGQKKG.PGI
 huCYP7     VLWASQANTI PATFWSLFQM IRNPEAMKAA TEEVKRTLEN AGQKVSLEGN
CYP11A1     MLAGGVDTTS MTLQWHLYEM ARNLKVQDML RAEV...LAA RHQAQGDM..
CYP11B1     LTAGSVDTTV FPLLMTLFEL ARNPNVQQAL RQES...LAA AASISEHP..
  CYP27     LLMAGVDTTS NTLTWALYHL SKDPEIQEAL HEEV...VGV VPAGQVPQ..
  CYP21     LLIGGTETTA NTLSWAVVFL LHHPEIQQRL QEELDHELGP GASSSRVP..
 CYP4A1     FMFEGHDTTA SGVSWIFYAL ATHPKHQQRC REEVQSVLGD GSSI...T..
 CYP1O1     LLVGGLDTVV NFLSFSMEFL AKSPEHRQEL IERPER.... ..........

                       ■          • •••                          385
  mHct1     SVHFTREQLD SLVCLESTIL EVLRLCSYSS IIREVQEDMN LSLESKSFSL
 huCYP7     PICLSQAELN DLPVLNSIIK ESLRLSSASL NIRTAKEDFT LHLEDGSYNI
CYP11A1     .....ATMLQ LVPLLKASIK ETLRLHPIS. VT..LQRYLV NDLVLRDYMI
CYP11B1     .....QKATT ELPLLRAALK ETLRLYPVG. LF..LERVAS SDLVLQNYHI
  CYP27     .....HKDFA HMPLLKAVLK ETLRLYPVV. PT..NSRIIE KEIEVDGFLF
  CYP21     .....YKDRA RLPLLNATIA EVLRLRPVVP LA..LPHRTT RPSSISGYDI
 CYP4A1     .....WDHLD QIPYTTMCIK EALRLYPPVP GI..VRELST SVTFPDGRSL
 CYO1O1     .......... ....IPAACE ELLRRFSLVA D....GRILT SDYEFHGVQL

                                  •          •                     434
  mHct1     RKGDFVALFP PLIHNDPEIF DAPKEFRFDR FI.EDGKKKS TFFKGGKRLK
 huCYP7     RKDSIIALYP QLMHLDPEIY PDPLTFKYDR YLDENGKTKT TFYCNGLKLK
CYP11A1     PAKTLVQVAI YALGREPTFF FDPENFDPTR WLSKDKNIT. .........Y
CYP11B1     PAGTLVRVFL YSLGRNPALF PRPERYNPQR WLDIKGSGR. .........N
  CYP27     PKNTQFVFCH YVVSRDPTAF SEPESFQPHR WLRNSQPATP RI.....QHP
  CYP21     PEGTVIIPNL QGAHLDETVW ERPHEFWPDR FLEPGKNSR. ..........
 CYP4A1     PKGIQVTLSI YGLHHNPKVW PNPEVFDPSR FAPDSPRHSH SF........
 CYP1O1     KKGDQILLPQ MLSGLDEREN ACPMHVDFSR .......... .........QK

                 ••o •    • •    •    •    o    o                  482
  mHct1     TYVMPFGLGT SKCPGRYFAV NEMKLLLIEL LTYFDLEIID ..RKPIGLNH
 huCYP7     YYYMPFGSGA TICPGRLFAI HEIKQFLILM LSYFELELIE GQAKCPPLDQ
CYP11A1     FRNLGFGWGV RQCLGRRIAE LEMTIFLINM LENFRVEIQH L.SDVGTIFN
CYP11B1     FYHVPFGFGM RQCLGRRLAE VEMLLLLHHV LKHLQVETLT Q.EDIKMVYS
  CYP27     FGSVPFGYGV RACLGRRIAE LEMQLLLARL IQKYKVVLAP ETGELKSVAR
  CYP21     ..ALAFGCGA RVCLGEPLAR LELFVVLTRL LQAF.....T LLPSGDALPS
 CYP4A1     ...LPFSGGA RNCIGKQFAM SEMKVIVALT LLRFELLPDP TKVPI.PLPR
 CYP1O1     VSHTTFGHGS HLCLGQHLAR REIIVILKEW LTRIP..... .DFSIAPGAQ

                       ■                  507
  mHct1     SRMFLGIQHP DSAVSFRYKA KSWRS..... .........
 huCYP7     SRAGLGILPP LNDIEFKYKF KHL....... .........
CYP11B1     L.....ILMP EKPISFTFWP FNQEATQQ.. .........
CYP11A1     F.....ILRP SMCPLLTFRA IN........ .........
  CYP27     I.....VLVP NKKVGLQFLQ RQC....... .........
  CYP21     L.....QPLP NCSVILKHQP FQVRLQPRGM GAHSPGQNQ
 CYP4A1     L.....VLKS KNGIYLYLKK LH........ .........
 CYP1O1     IQHKSGIVSG VQALPLVWDP ATTKAV.... .........
```

unique to mHct1 and CYP7 appear at positions 334-340, 368, 369, and 427-431.

Figure 5.4 indicates three amino acids (residues 84, 136 and 294) which appear to be generally conserved amongst CYP proteins except mHct1 and CYP7. In addition, there is described a conserved glycine at amino acid position 291 which is always followed by a threonine three amino acids towards the C-terminus at position 294. This domain (G/AGXXT) is proposed to be the O_2-binding pocket of cytochromes P-450 (Poulos, 1988), however this structure is not conserved in CYP7 or mHct1. The glycine is not conserved at position 291, both CYP7 and mHct1 having a serine. There is however a threonine in both proteins at position 295, and an alanine precedes the position where the glycine is not conserved. The motif in CYP7 and mHct1 may therefore be ASXXNT, or L/QXXT. The same situation is also found in CYP11B1, however there is a glycine at position 290 (AGXXXT). Nevertheless, it appears that CYP7, mHct1, and possibly also CYP11B1 have a slightly modified structure at the O_2-binding pocket. This is supported by the high degree of identity in this region (amino acids 288-301) between CYP7 and mHct1 (see Figure 5.4).

The N-terminus of the Hct1 protein is hydrophobic, which is typical of microsomal cytochromes P-450. It is this portion of other cytochromes P-450 which has been shown to insert into the membrane of the endoplasmic reticulum, holding the main bulk of the protein on the cytoplasmic side (reviewed in Kemper and Szczesna-Skorupa, 1989). Consistent with microsomal cytochromes P-450, the N-terminus lacks basic amino acids prior to a hydrophobic core (amino acids 9 - 35). Chimaeric proteins which have N-terminal sequences fused to non-microsomal proteins have shown that 20 - 30 amino acids from the N-terminus are sufficent for association with microsomal membranes (Kemper and Szczesna-Skopura, 1989 and references therein).

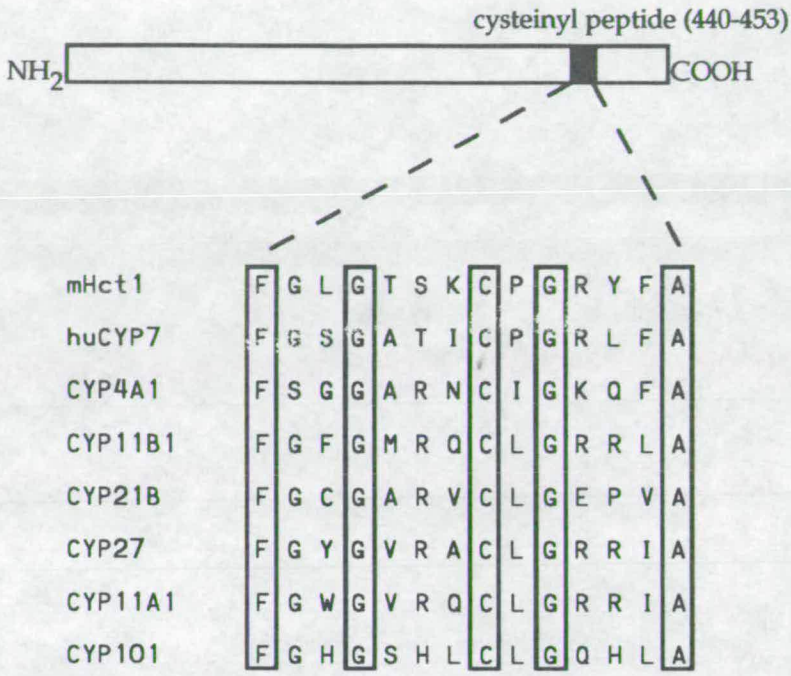
Virtually all cytochrome P-450 sequences have a strictly conserved motif, FXXGXXXCXG (Figure 5.5(a)) (amino acids 440-449 in Figure 5.4), which is present in 202 of the known 205 sequences (Nelson *et al.*, 1993). In the three exceptions, the changes involve the non-cysteine

Figure 5.5(a) Conservation of the haem-binding domain in mouse Hct1 protein.

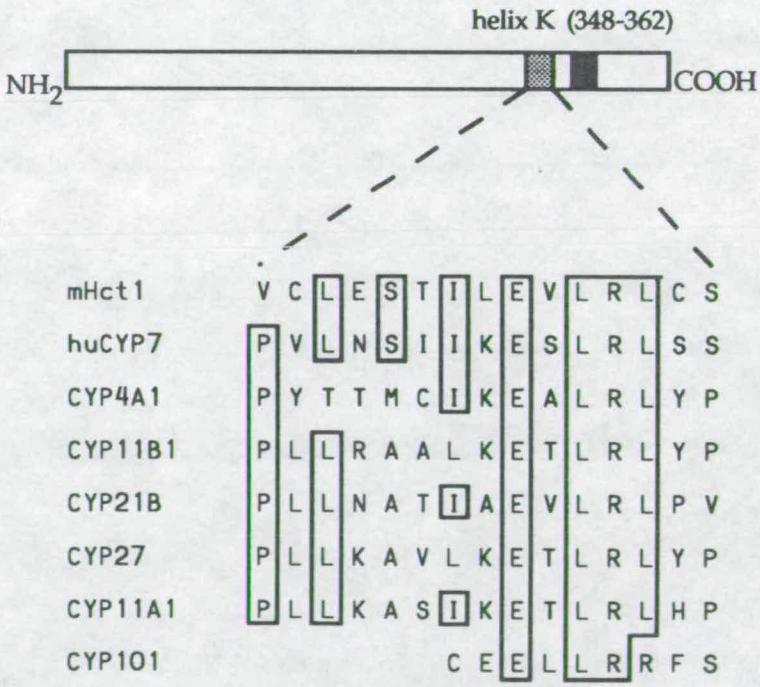
Mouse Hct1 was aligned with the haem-binding regions of CYP7, human cholesterol 7α-hydroxylase (Noshiro and Okuda, 1990), CYP4A1, rat lauric acid ω-hydroxylase (Hardwick *et al.*, 1987), CYP11B1, human steroid 11β-hydroxylase (Mornet *et al.*, 1989), CYP21B, human steroid 21-hydroxylase (sequence from Nelson and Strobel, 1988), CYP27, human sterol 27-hydroxylase (Cali and Russell, 1991), CYP11A1, human P-450scc (cholesterol side chain-cleavage enzyme) (Morohashi *et al.*, 1987), and CYP101, bacterial cytochrome P-450cam (Haniu *et al.*, 1982). The alignment was produced using the UWGCG program PILEUP with all parameters set at default values.

(b) Alignment of mouse Hct1 helix K (conserved steroidogenic domain).

(a)


cysteinyl peptide (440-453)
NH2
COOH
mHct1 F G L G T S K C P G R Y F A
huCYP7 F G S G A T I C P G R L F A
CYP4A1 F S G G A R N C I G K Q F A
CYP11B1 F G F G M R Q C L G R R L A
CYP21B F G C G A R V C L G E P V A
CYP27 F G Y G V R A C L G R R I A
CYP11A1 F G W G V R Q C L G R R I A
CYP101 F G H G S H L C L G Q H L A

(b)


helix K (348-362)
NH2
COOH
mHct1 V C L E S T I L E V L R L C S
huCYP7 P V L N S I I K E S L R L S S
CYP4A1 P Y T T M C I K E A L R L Y P
CYP11B1 P L L R A A L K E T L R L Y P
CYP21B P L L N A T I A E V L R L P V
CYP27 P L L K A V L K E T L R L Y P
CYP11A1 P L L K A S I K E T L R L H P
CYP101 C E E L L R R F S

residues. Thus the cysteine residue is present in all cytochromes P-450, and is generally located towards the carboxy terminus. In the case of mHct1, this cysteinyl peptide is located at amino acids 440-449. The cysteine is believed to be involved in the binding of the haem cofactor. The arrangement of amino acids surrounding the cysteine residue preserves the three-dimensional structure of this region for ligand binding (Poulos, 1988).

A domain which lies proximal to the haem-binding domain, referred to as helix K (Poulos, 1988) or the conserved steroidogenic domain by Noshiro and Okuda (1990), also shows conservation, in particular, with those cytochromes P-450 which are involved in steroidogenic conversions (Figure 5.5(b) (amino acids 352-366 in Figure 5.4). mHct1 shares a number of conserved amino acids, however once again, there is a non-conserved change from proline to valine at position 340.

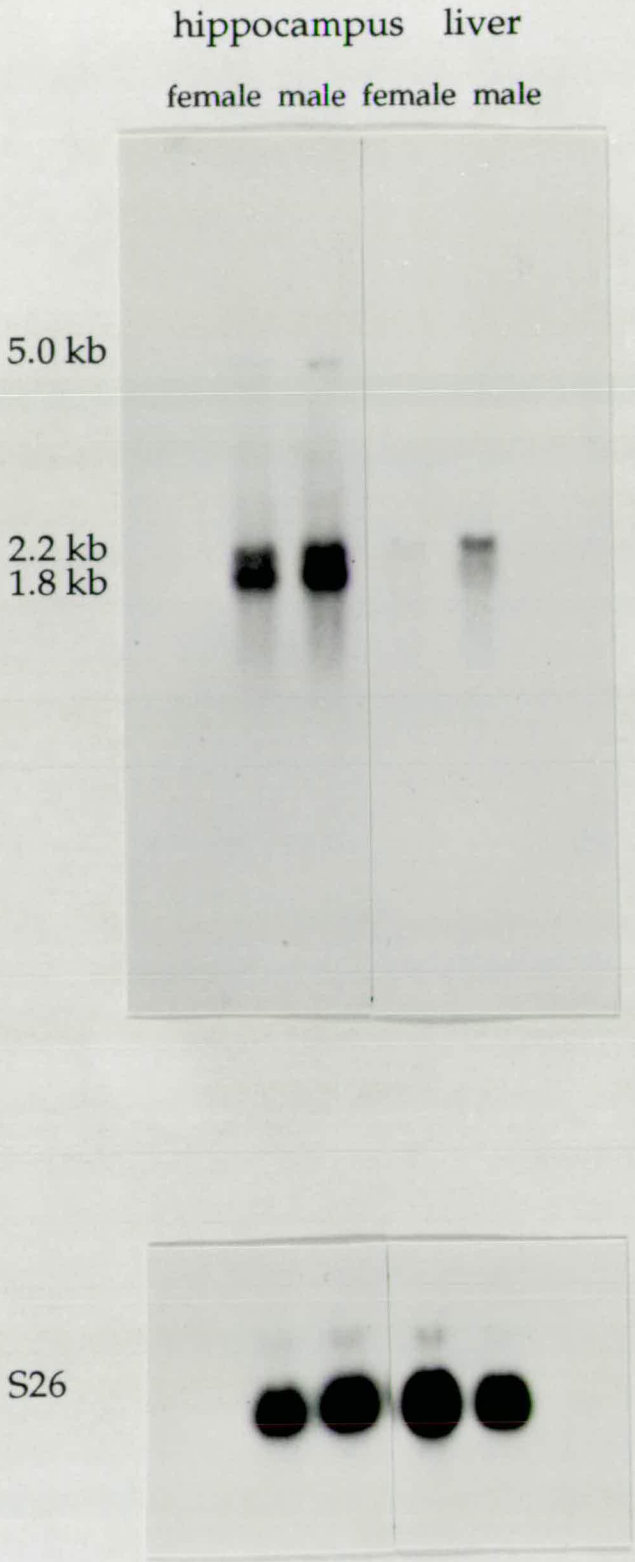
5.2 Sexually dimorphic regulation of Hct1 in the rat

A number of cytochromes P-450 are expressed in the liver in a sexually dimorphic pattern (for review, see Morgan and Gustaffson, 1987). In addition, Roof and Havens (1992) have observed a sexual dimorphism in the size of the granule cell layer of the dentate gyrus. Sexually dimorphic gene expression of Hct1 was examined in the rat liver and hippocampus. The 1.4 kb Hct1 clone was labelled with ^{32}P by random primer labelling, and was used to probe a Northern blot of RNA samples extracted from the liver and hippocampus of female and male adult rats (Fig. 5.6).

No changes in Hct1 expression levels were observed in the male and female hippocampus. However, Hct1 expression in the liver was found to be sexually dimorphic (Figure 5.6). Hct1 is expressed in the male liver, while Hct1 mRNA in the female liver was detected by Northern blot analysis only after prolonged exposure.

Figure 5.6 Sexually dimorphic expression of Hct1 in the rat.

Hippocampus RNA samples were polyA$^+$-selected, while liver samples were prepared as total cytoplasmic RNA. 1 μg of polyA$^+$ samples and 10 μg of total cytoplasmic RNA was used for Northern analysis. Hybridisation to ribosomal protein S26 was used to determine equivalent loading of samples.


hippocampus
liver
female
male
female
male
5.0 kb
2.2 kb
1.8 kb
S26

5.3 Expression of Hct1 protein for enzymatic activity analysis

Sequence comparisons predicted an enzymatic activity for mHct1, but not a substrate preference. Thus, to confirm catalytic activity and substrate specificity, a yeast heterologous expression system was used to produce recombinant Hct1 protein. Because clones 35 and 40 of mHct1 varied in 5' and 3' non-coding sequences (see section 4.3 and Figure 4.3) which could possibly affect Hct1 expression, both cDNA clones were used in the expression experiments.

The pYES II expression vector and the *S. cerevisiae* strain, INVSc1, were provided by Dr. J. Tsang, Dundee. mHct1 clones were inserted into the vector using restriction sites in the polylinker. The correct orientation of the inserts was checked using unique SacI and BamHI sites contained within the cDNA (refer to Figure 4.3). The constructs, containing mHct1 under the control of the inducible promoter, Gal1, were transformed into INVScI. pYES II plasmids contain the URA3 gene, allowing selection of transformants on minimal medium without uracil. A plasmid containing no mHct1 cDNA insert was also transformed into yeast.

The recombinant yeast were grown to late logarithmic phase with appropriate selection in minimal medium containing 2% glucose. To induce expression, the medium was changed, replacing the glucose with 2% galactose. A time course of samples was taken to assay for mHct1 expression.

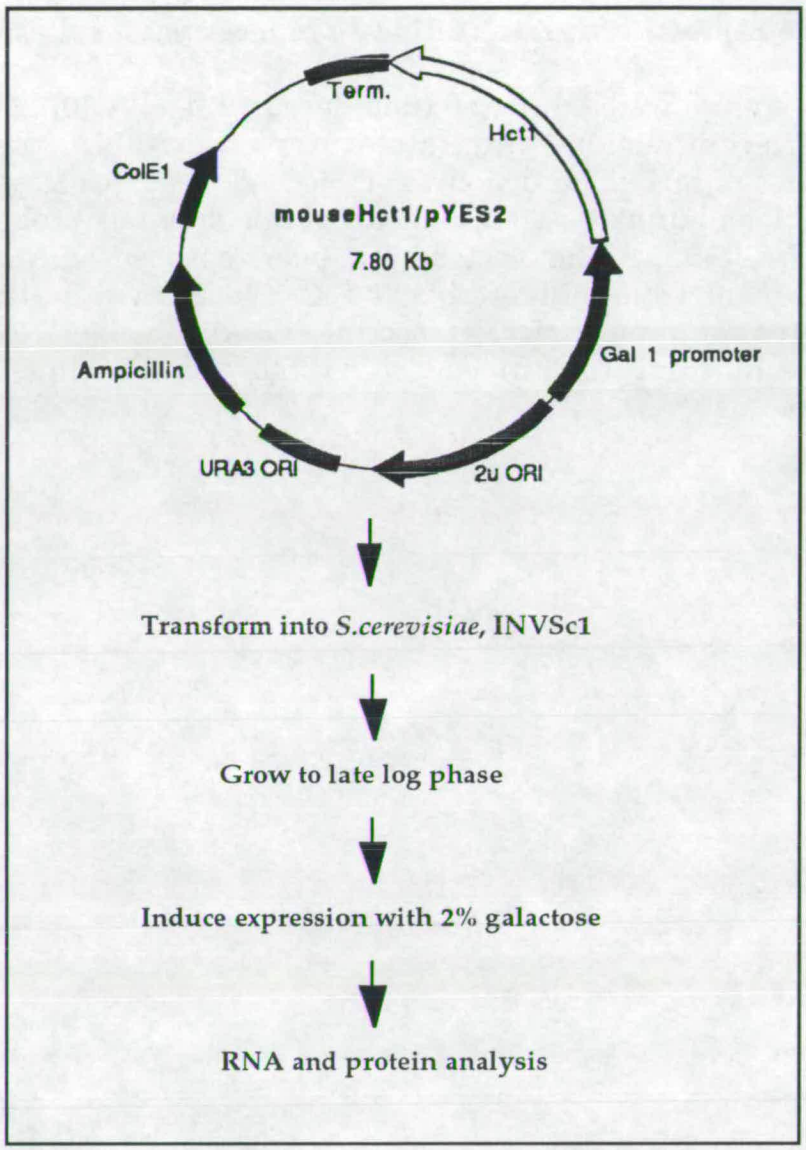
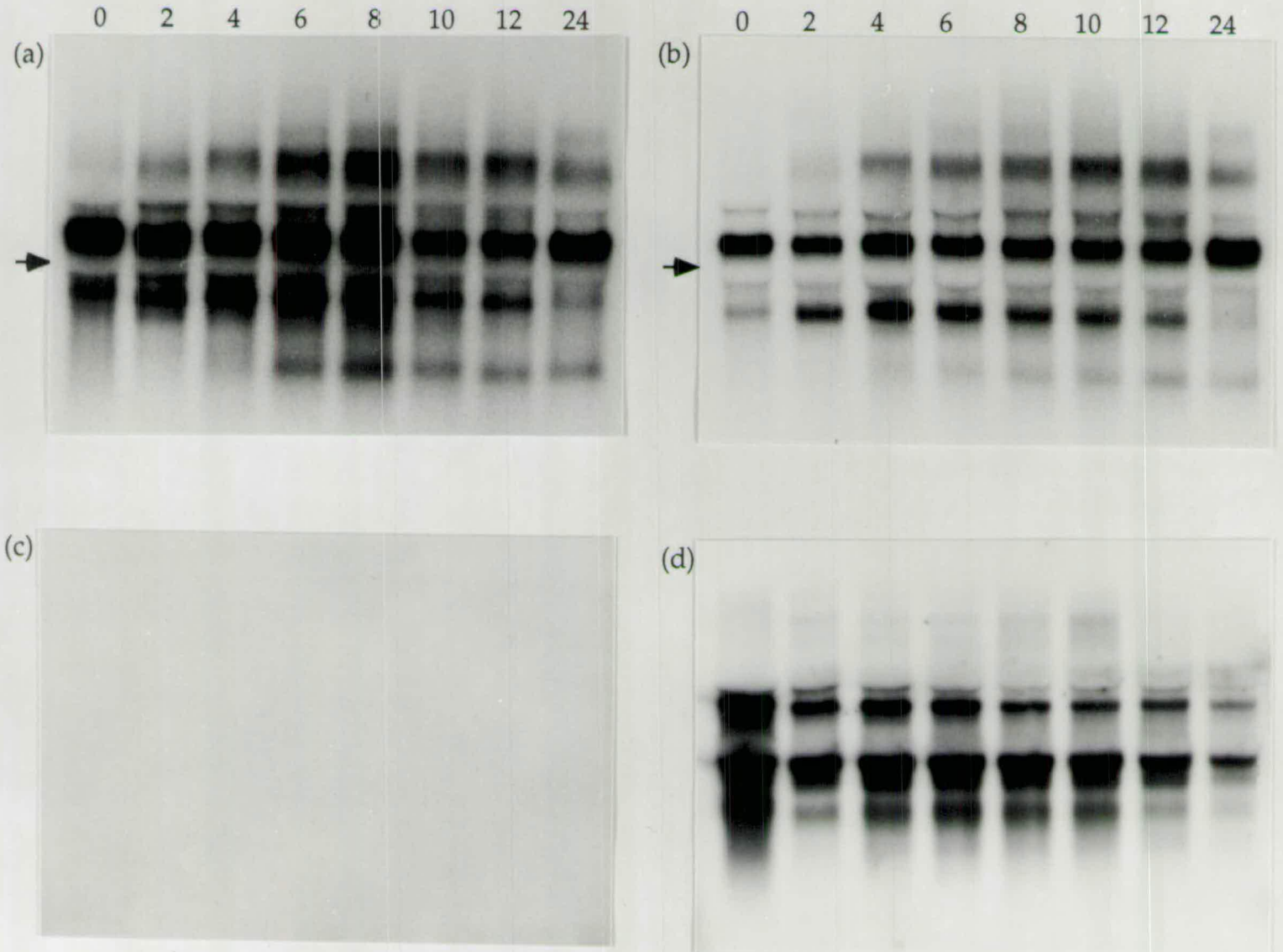


Figure 5.7 Expression vector for Hct1 production in yeast.
The full length cDNA of Hct1, clone 35 from the mouse, was inserted into the pYES II expression vector using a NotI site in the polylinker located between the Gal1 promoter and the termination signal. Correct orientation was determined by restriction enzyme analysis. The plasmid was transformed into yeast as described in Section 7.2.8. To induce Hct-1 expression, cultures were grown to late log phase in minimal media with uracil selection and glucose as a carbon source. Expression of Hct1 was induced by replacing glucose with galactose.

Northern blot analysis was used to examine Hct1 mRNA production. Total cytoplasmic RNA was extracted from culture samples

Figure 5.8 Expression of mHct1 RNA in induced yeast cultures.

Culture samples were taken at 0 (uninduced), 2, 4, 6, 8, 10, 12, and 24 hours after induction, and total cytoplasmic RNA prepared. Timecourse of mHct1 induction: (a) mHct1 (clone 35) expression, (b) mHct1 (clone 40) expression, and (c) vector only, all probed with mHct1 clone 35. (d) The 'vector only' filter reprobed with a 1.3 kb BglI fragment from pBluescript II SK. The arrows indicate the position of the mouse Hct1 transcript (1.8 kb) in liver RNA (the exposure time was insufficient to visualise endogenous mHct1 message).


0 2 4 6 8 10 12 24
(a)
0 2 4 6 8 10 12 24
(b)
(c)
(d)

taken at each of the induced timepoints and the uninduced sample. Each RNA sample was fractionated through a denaturing agarose gel, transferred to a Hybond membrane by Northern blotting, then probed with the full length mouse Hct1 cDNA labelled with ^{32}P by random primer labelling (Figure 5.8).

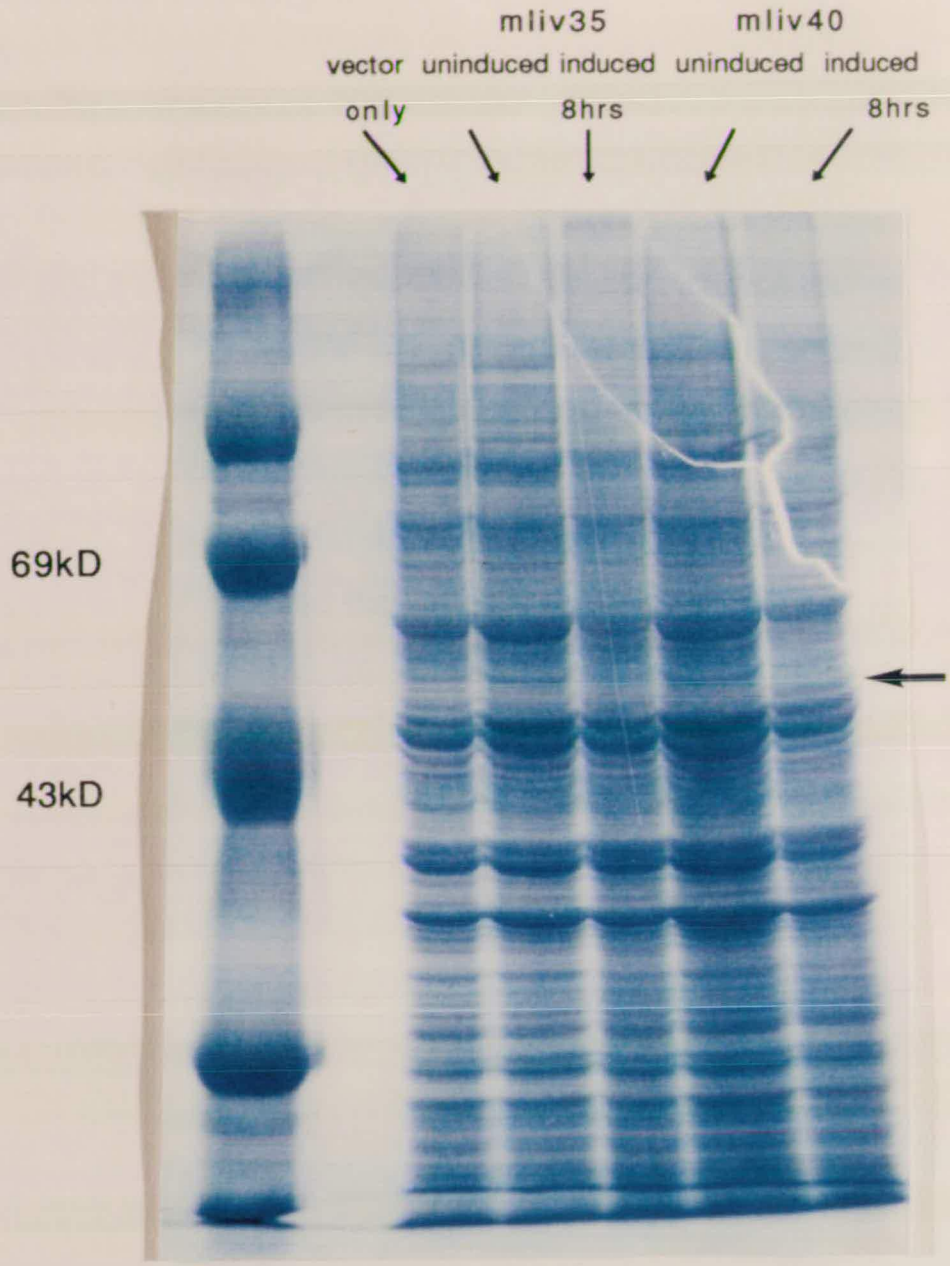
The mouse Hct1 cDNA probe hybridised to both induced and uninduced samples for both clone 35 and clone 40 (Figure 5.8(a) (b)). This is not due to cross-hybridisation to endogenous yeast transcripts, as mHct1 does not hybridise to the 'vector only' samples (Figure 5.8(c)). In addition, there was no apparent increase in transcripts of the expected size of 1.8 kb. There were transcripts of 0.56 kb, 1.1 kb and 6.6 kb which did increase with galactose induction, however it was not clear whether these could correspond to transcribed Hct1.

The possibility that the range of bands observed corresponded to contaminating plasmid was tested. The pYes II plasmid is a high copy number plasmid, and although the RNA preparation method should exclude DNA contamination, even minute amounts of plasmid contamination could result in the observed hybridisation pattern. A 1.3 kb piece of pBluescript II was generated by BglI digestion, and was used to reprobe the 'vector only' filter (Figure 5.8(d)). The same pattern of hybridisation was found using a pBluescript II probe, thus it was concluded that the bulk of the signal was contributed by plasmid contamination. The 0.56 kb 'induced' band was not recognised by the vector probe, while the 1.1 kb and 6.6 kb 'induced' bands were recognised by the plasmid probe. However the results from the RNA analysis were inconclusive, and because ultimately Hct1 protein production must be demonstrated, RNA analysis was not pursued further.

To detect Hct1 protein production, total protein extracts were prepared from uninduced cultures and cultures after 8 hours of induction. In the absence of an antibody specific for Hct1, 30 µg of each protein extract was fractionated on a 10% SDS-PAGE gel, and then stained using Coomassie brilliant blue (Figure 5.9).

The products of most cytochrome P-450 genes, like their transcript sizes, tend to be of similar size, with protein molecular weights

Figure 5.9 Expression of Hct1 protein in yeast after an 8 hour induction. 30 μg of total protein extract was fractionated through a 10% SDS-PAGE gel, followed by staining with Coomassie brilliant blue.


mliv35
mliv40
vector only
uninduced
induced 8hrs
uninduced
induced 8hrs
69kD
43kD

between 49 kD and 55 kD (Paine, 1991). Thus Hct1 protein in the induced samples was predicted to be in this range. A band in the correct size range was present in the induced sample of clone 40, and absent in the uninduced sample, as indicated by the arrow in Figure 5.8. An induced band of weaker intensity can also be seen in the clone 35 sample at the same position, however this is concealed somewhat as the uninduced sample is overloaded.

5.4 Discussion

The Hct1 gene is present in rat, mouse and human, and appears to be a single copy gene. While cytochrome P-450 genes are scattered over both mouse and human chromosomes, cytochrome P-450 subfamilies tend to be clustered together on the same chromosome. The human *CYP2A* and *2B* subfamilies are linked to chromosome 19, *CYP2C* and *2E* subfamilies are located on human chromosome 10, and the mouse *cyp2a*, *2b*, and *2e* subfamilies are linked to mouse chromosome 7 (reviewed by Paine, 1991). Human cholesterol 7α-hydroxylase (*CYP7*) is located on chromosome 8q11-q12 (Cohen *et al*, 1992). It will therefore be of future interest to examine the chromosomal location of the human orthologue of Hct1 as a measure of linkage to the *CYP7* family. In addition, should Hct1 map to a known genetic locus, this may suggest a function for Hct1 in human.

The structure of mHct1 and its similarity to other cytochromes P-450 suggests a possible function for Hct1. Hct1 is most homologous to cholesterol 7α-hydroxylase, both of which are more similar to each other than any other cytochrome P-450. The striking conservation of amino acid sequence between the two proteins (described section 5.2) suggests that CYP7 and Hct1 are most likely members of the same cytochrome P-450 family, and that these two enzymes have similar activities. Based on the Hct1 leader sequence, Hct1 most likely resides in the endoplasmic reticulum and not in the mitochondria, the other principal cellular site of cytochrome P-450 activity. In general, enzymes that form bile acids from cholesterol can be divided into two groups: those which catalyse oxidation of the cholesterol side chain tend to be

located in the mitochondria, like CYP11A1 (Morohashi *et al*, 1987), while those enzymes like cholesterol 7α-hydroxylase which modify the ring structure of cholesterol are located on the surface of the endoplasmic reticulum (Bjorkhem, 1985). Considering the similarities between CYP7 and Hct1, it might therefore be suggested that Hct1 enzymatic activity is directed to the cholesterol ring structure.

The domain which is most highly conserved in steroidogenic cytochromes P-450 is also conserved in mHct1 (Figure 5.5(b)). This once again supports a possible role for mHct1 in the modification of cholesterol-based molecules. This helix K region is thought to participate in binding substrate (Li *et al*, 1990). However the non-conserved proline at position 340 may indicate a different substrate interaction with mHct1.

The structure of the O_2-binding pocket of mHct1 and CYP7 is not conserved as proposed by Poulos (1988). Crystallographic studies of the bacterial CYP101 indicate that the corresponding threonine at position 252 disrupts the helix formation in that region, and is important in providing a structural pocket for an oxygen molecule (Poulos *et al*, 1987). In addition, Imai and Nakamura (1989) have shown by site-directed mutagenesis of the conserved threonine in both CYP4A1 and CYP2C11 that this region can also influence substrate specificity and affinity. This modified region in CYP7 and mHct1 suggests that these proteins are structurally different on the distal side of the haem, and this may be reflected both in a modified oxygen interaction and in substrate choice.

While Hct1 hippocampus expression is equivalent in both male and female rats, Hct1 expression in male rat liver is not unlike other cytochromes P-450. *CYP2C12* is expressed preferentially in the liver of the female rat, while *CYP2C11* is expressed similarly to Hct1 in male rat liver (Donahue *et al*, 1991). This expression pattern for the *CYP2C* family members has been shown to be regulated by the sexually dimorphic growth hormone secretory pattern (Morgan and Gustaffson, 1987). It is possible that Hct1 is under the same regulatory controls which govern *CYP2C11* expression but, because Hct1 is so distantly related to this family, it is possible that separate regulatory mechanisms

exist. Whatever the mechanism, Hct1 expression in the hippocampus appears to be excluded from such regulation.

Sexually dimorphic Hct1 expression was only examined in the rat, and it cannot be assumed that the same regulation of Hct1 expression exists in other species. Indeed, it has been shown by Noshiro *et al* (1988) that sexually dimorphic expression is not always conserved even within mouse strains. Nevertheless, Hct1 expression in the rat liver is sexually dimorphic. The consequences of such gene expression may become apparent once a specific role for Hct1 in the liver is defined.

CHAPTER SIX

DISCUSSION

Hct1 was identified during a differential screen designed to isolate rat cDNA clones corresponding to transcripts enriched or uniquely expressed in the hippocampus. Within the brain, rat Hct1 is expressed predominantly in the hippocampus, with very low levels of expression detectable in cortex and other brain areas. In addition, low level expression of rHct1 is detected in liver and kidney but not in any other tissues tested. In contrast to rat, the expression pattern of the mouse orthologue is widespread in adult brain, showing no compelling enrichment in hippocampus.

A full-length mouse cDNA of Hct1 shows significant nucleic acid and amino acid homologies to cholesterol 7α-hydroxylase (CYP7), an enzyme required for the conversion of cholesterol to bile acids. CYP7 belongs to the cytochrome P-450 superfamily of monooxygenases, a large group of enzymes involved in diverse processes including steroidogenesis and metabolism of exogenous compounds. Based on sequence homology, it is likely that Hct1 constitutes the second member of this cytochrome P-450 family.

Hct1 was cloned originally to assess its suitability for use as a promoter to drive hippocampus-specific transgene expression. The potential use of Hct1 regulatory elements in transgenesis and the function of Hct1 as a cytochrome P-450 in the hippocampus remain to be determined. The application of Hct promoters, and possible roles and substrates for Hct1 are discussed.

6.1 The application of Hct promoters in transgenesis

The use of promoters which can direct transgene expression specifically to the hippocampus would provide a means of investigating the effects of modified gene expression in the

hippocampus, without affecting other areas of the animal. Furthermore, the use of tissue-specific gene targeting in mice further necessitates the isolation of hippocampus-specific promoters.

As discussed previously (section 1.4.2(a)), the consequences of a developmental role for gene products such as Fyn can complicate analyses of possible adult phenotypes. An Hct gene promoter may help to distinguish between developmental effects and a specific role for Fyn in spatial learning. Directed expression of Fyn to the hippocampus of the null mouse should "rescue" the learning deficit if Fyn plays a specific role in learning. Similarly, LTP should be restored in "Fyn-rescued mice" if Fyn is an essential component of signalling in LTP. This rationale can be applied also to the CaMKII and PKCγ null mice, where any non-specific effects of gene inactivation in other cell types can be eliminated with a hippocampus-specific rescue of the learning deficit.

The use of the Cre/*lox*P system to inactivate gene expression or to introduce gene variants at a specific location (see section 1.4.2(c)), expands the potential of transgenic analysis in the study of hippocampus function. Specifically, the use of hippocampus-targeted gene inactivation can facilitate the analysis of protein function where a conventional knockout results in lethality. For example, the NMDAR1 subunit knockout mouse (Li *et al.*, 1994), if inactivated in the hippocampus alone, may result in viable offspring allowing analysis of an adult hippocampus phenotype.

Another example of the potential application of Hct promoters would be to target subtle modifications of a gene specifically to the hippocampus, rather than generating a null allele. Again with the example of the NMDAR1 receptor subunit, subtle changes to the subunit could be made by altering the endogenous gene in the hippocampus. For instance, the role of phosphorylation of the NMDA receptor subunits in LTP and subsequent behavioural effects could be assessed in the animal by mutation of one or all phosphorylation sites. Mice could be generated which have subtle changes in all of the receptor subunits of the NMDA receptor, which, upon breeding, could allow the analysis of combinations of modified NMDA receptor

subunits. It should, however, be re-emphasised that tissue-specific gene inactivation, such as the preceding examples, is still limited by the lack of hippocampus-specific promoters.

It is worth noting that while mouse transgenesis is of great potential, the rat is the preferred experimental model for behavioural analysis. The use of mice for behavioural analyses is at an early stage and extrapolating data from the rat system directly to mice may not be completely valid. It may therefore be necessary to repeat the hippocampus ablation studies in the mouse to confirm the legitimacy of the mouse system, taking into account genetic background. Studies have shown that different strains of mice exhibit varying abilities in spatial learning in the water maze. For example, C57 mice performed better than DBA mice in one study (Upchurch and Wehner, 1988). In addition, the water maze spatial learning test, used to assess rat hippocampus-dependent learning, may need adaptation to accommodate mouse analysis. Thus the size of the pool and the age of the animals used may need to be considered when assessing the ability of the mouse to perform the task. Finally, patterns of gene expression may not necessarily be conserved between species. Such differences in gene expression between mouse and rat, as seen with Hct1, again support the idea that mouse and rat are two unique experimental systems that do not necessarily overlap.

6.2 The suitability of Hct1 regulatory elements for transgenesis

The use of Hct1 regulatory elements to direct expression to the hippocampus is currently restricted to rat transgenesis or the use of the rat gene promoter in the mouse. As described in section 3.3, mouse Hct1 is expressed throughout the brain, showing no enrichment in the hippocampus. Transgenic analysis of Hct1 function in the hippocampus therefore would also be restricted to addition transgenesis, as gene inactivation using homologous recombination is not yet possible in rat.

Rat Hct1 is enriched in the hippocampus, however very low level expression is found outside the hippocampus and in extraneuronal

tissue. Thus Hct1 promoter elements are unlikely to be capable of directing gene expression exclusively to the hippocampus. The use of the Hct1 promoter, however, may still be possible in some types of experiments where transgene expression in extrahippocampal tissue and at very low levels would be unlikely to lead to ambiguities of interpretation. Ectopic expression of molecules such as the NMDA receptor subunits in liver, for example, may not bear any consequence on behavioural analyses and would be unlikely to affect the electrophysiology of the hippocampus. In addition, it may be possible to use the Hct1 promoter when relatively high levels of transgene expression are required. Expression outwith the hippocampus of a transgene modified to exert a dominant-negative effect, would not be expected to compete with the endogenous protein if expressed at very low levels.

An alternative route to exploiting the Hct1 promoter would be to determine which elements in the promoter are responsible for the spatial regulation of expression (described further in section 6.3). Deletion analyses, as described in section 6.3, may identify regions within the promoter specifically responsible for Hct1 expression in liver or hippocampus. Thus, through such analyses, the possibility exists to create artificial hippocampus-specific promoters from Hct1.

6.3 Regulation of gene expression in the hippocampus

Gene transcription by RNA polymerase II is controlled by the interaction of *trans*-acting proteins at specific DNA elements in the promoter or enhancer. It is this assembly of tissue-specific and ubiquitous transcription factors together with the basal transcription machinery that direct the initiation of transcription. Promoter deletion studies and protein-DNA binding studies have identified DNA elements required for tissue-specific expression, and the transcription factor which bind to such elements. The identification of promoter elements and transcription factors which control hippocampus expression would be of potential utility.

Purkinje cell protein 2 (Pcp-2) is expressed in cerebellar Purkinje cells and retinal bipolar neurons. The regulatory elements which direct

expression to these two cell types were dissected using deleted promoter constructs to drive expression of a *lacZ* reporter gene in transgenic mice (Vandaele *et al.*, 1991). A DNA element conferring neuronal expression was identified and was shown to be present in a number of other promoters of neuronal genes. In addition, cerebellar expression could be separated from retinal expression. Similarly, genes expressed specifically in neurons of the olfactory epithelium all contain DNA binding sites for the olfactory-specific transcription factor, Olf-1 (Wang *et al.*, 1993).

The analysis of Hct promoters may similarly reveal consensus sites for transcription factors required for hippocampus expression. A very simple approach would be to examine promoter sequences for conservation of sequence motifs. A more thorough analysis would involve promoter deletion constructs and examination of reporter gene expression in transgenic mice. This can be followed by analysis of *in vitro* protein-DNA interactions (for example see Somma *et al.*, 1991) and finally by *in vivo* analysis of protein interactions at the promoter (Mueller and Wold, 1989). Such strategies could be applied, for example, to the serotonin receptor where expression is restricted to two distinct regions in the brain (Wisden *et al.*, 1993). Transgenic approaches, like those used for Pcp-2 by Vandaele *et al.* (1991), may discriminate between DNA regions directing gene expression in the hippocampus or in the habenulae. By this approach, it is possible that artificial Hct promoters could be created by removing the promoter region required for expression in the habenulae.

6.4 Cytochrome P-450 gene expression in the brain

The exact role of cytochrome P-450 activity in brain is not defined, however the detection of multiple forms of cytochrome P-450 in brain suggests a variety of functions. A summary of reported expression or activity of cytochromes P-450 in the central nervous system is described in Appendix I. The emerging pattern of expression appears to be largely restricted to members of CYP1, CYP2, and CYP3 families, proteins involved largely in foreign compound metabolism and procarcinogen activation, and CYP4 family members which hydroxylate fatty acids

such as arachidonic acid. The CYP2 family which is involved in the metabolism of foreign compounds are expressed widely throughout the brain, and are induced upon different treatments such as ethanol and nicotine administration (Warner *et al.*, 1994; Anandatheerthavarada *et al.*, 1993). The existence of these CYP proteins in the brain suggests a capacity for the brain to metabolise foreign compounds. This is supported by the upregulation of CYP1A, CYP2C and CYP4A in brain on partial hepatectomy (Warner *et al.*, 1993). P-450-dependent metabolism is thus of particular relevance in the design of neuro-specific drugs which may affect a drugs action as well as the potentially deleterious effects of the resulting metabolites.

Interestingly, Stromstedt *et al.* (1993) have detected 6-alpha and 7-alpha hydroxylation activities on the anaesthetic steroid 3α-hydroxy-5α-pregnan-20-one, which was shown to be cytochrome P-450-mediated. The expression of CYP7 in brain is not known, however these hydroxylation activities could be mediated by CYP7 and perhaps Hct1. Finally, expression of cytochromes P-450 is not restricted to neurons as CYP2E1 is expressed in glial cells of the cerebellum (Hansson *et al.*, 1990) and expression of several CYP genes has been shown in the rat glioma cell line C6 (Geng and Strobel, 1993). Thus, both sub-populations of neurons and glia possess functional cytochrome P-450 systems.

6.5 A role for Hct1 in the hippocampus

Until the preferred substrate for Hct1 is identified, prediction of a role for Hct1 in the hippocampus can only be speculative. However, a suggested role for Hct1 in cholesterol homeostasis is not unreasonable, given the significant sequence homology between Hct1 and CYP7. Cholesterol is a major constituent of plasma membranes and changes in membrane structure might be predicted as a result of cholesterol mobilisation. It should again be noted, as discussed in section 1.3.4(a)), that nucleic acids encoding the enzyme responsible for the 12α-hydroxylation of cholesterol in the bile acid biosynthetic pathway have not as yet been cloned. Hct1 must be considered a candidate for this

enzymatic activity in the liver, but would presumably act in another context in the hippocampus.

In liver, nitric oxide (NO) has been shown to inhibit the activity of CYP enzymes (Stadler *et al.*, 1994). This occurs at the level of transcription, by an unknown mechanism, as well as post-translationally, by binding to the haem cofactor so as to inhibit catalytic activity. Although the subcellular distribution of Hct1 in neurons is not known it might be predicted that, if Hct1 is postsynaptic, cytochrome P-450 activity would be down-regulated by NO produced during LTP. The consequences of possible Hct1-NO interactions at the postsynapse may be related, for example, to downregulation of cholesterol mobilisation from the plasma membrane. It should be noted that carbon monoxide, another putative retrograde messenger, would also be expected to bind to the haem group and disrupt cytochrome P-450 activity.

Arachidonic acid might also be considered a candidate substrate for Hct1, although arachidonic acid is a known substrate for members of the CYP4 family of cytochromes P-450 (section 1.3.4(a)), and Hct1 shows only weak homology to this family. However the pathways of arachidonic acid metabolism are not yet fully elucidated, and it has been proposed that other cytochrome P-450-like activity may be involved in some of these pathways (Piomelli and Greengard, 1990). For example, the enzymes which participate in the pathway which produces arachidonylethanolamide (also known as "anandamide") are not characterised. Anandamide was identified as a possible endogenous ligand for the cannabinoid receptor (Devane *et al.*, 1992) and is most likely a product of arachidonic acid metabolism.

It is also possible that a P-450-derived metabolite of arachidonic acid is responsible for the effects attributed to arachidonic acid itself. For example, it will be of interest to investigate the role of possible Hct1 metabolites as retrograde messengers in long-term potentiation; so far, only cyclooxygenase and lipoxygenase derivatives of arachidonic acid have been addressed (Williams *et al.*, 1989a).

6.6 A role for Hct1 in neurodysfunction?

Associations have been observed between cholesterol and a number of human neurodegenerative diseases. Thus proteins which are involved in the regulation of cholesterol levels may also have a role in processes leading to neurodegeneration. Parkinson's disease, Adrenoleukodystrophy, Type C Niemann-Pick disease and Alzheimer's disease are examples of neurodegenerative disease associated with changes in cholesterol homeostasis.

Wilson and Sargent (1993) have noted a 25% increase in cholesteryl ester levels in the brain of patients with adrenoleukodystrophy, a fatal disease resulting from the progressive demyelination of the CNS and PNS, as well as adrenal insufficiency. In contrast, a reduced biosynthetic cholesterol capability has been shown in fibroblasts from patients with Parkinson's disease (Musanti *et al.*, 1993). In neurological deterioration characteristic of Niemann-Pick disease type C there is a defect in intracellular cholesterol transport. Exogenous cholesterol from lysosomes, the principal site of cholesterol storage, is not transported for processing, resulting in an accumulation of cholesterol in the cell (Vanier *et al.*, 1991).

A possible link between Alzheimer's disease and cholesterol metabolism was described when the frequency of a particular allele of the apolipoprotein E gene was found to be elevated in patients with late onset Alzheimer's disease (Corder *et al.*, 1993). ApoE is a component of plasma lipoproteins which are responsible for the transport and clearance of cholesterol. The ApoE type 4 allele is found in 80% of familial late onset Alzheimer's disease while in control patients the E3 allele predominates. The E4 allele has a single amino acid change from the most common E3 allele. The effects of carrying one or two type 4 alleles and an association with Alzheimer's disease is not known, however ApoE4 is a component of amyloid plaques found in Alzheimer's brains. Further, the amyloid (the βA4 peptide) binds to ApoE with high avidity (Strittmatter *et al.*, 1993). Thus ApoE is implicated in the pathogenesis of Alzheimer's disease, with the suggestion of resulting effects on cholesterol homeostasis.

Alzheimer's disease is associated with the deposition of amyloid (βA4 peptide) produced by enzymatic cleavage of the amyloid precursor protein gene (APP). APP is a membrane-spanning protein, however the cleavage point on APP to produce the βA4 peptide is within the membrane. It is thus not clear how the enzyme responsible for the cleavage might access the cleavage site. One interesting hypothesis is that Alzheimer's pathology results from an increased mobilisation of cholesterol from the plasma membrane, causing structural perturbation of the membrane. Mason *et al.* (1992) have shown a decrease in the thickness of the lipid bilayer of the plasma membrane, and a 30% decrease in cholesterol content in plasma membranes from Alzheimer's patients. It was proposed that this might result in exposure of the cleavage site on APP, allowing for the abnormal production of the βA4 peptide. Based on such hypotheses it could be suggested that an enzyme such as Hct1, which is expressed in one of the preferred sites of amyloid deposition and which may be involved in cholesterol metabolism, could play a role in Alzheimer's disease. A gene for Hct1 has been detected by Southern blot analysis in human, and the preceding hypothesis assumes that human Hct1 is expressed in the hippocampus. Experiments are however underway to test this prediction by genetic mapping of the Hct1 orthologue in human.

6.7 The future for Hct1

The normal function of Hct1 and the suitability of the Hct1 promoter for use in hippocampus transgenics will be the focus of future experiments.

Immediate experiments will focus on the substrate preference for Hct1. This involves addition of a candidate substrate to a microsomal preparation from the recombinant yeast expressing mouse Hct1. The substrate, appropriately labelled, can be assayed for modification by thin layer chromatography (TLC) or high performance liquid chromatography (HPLC). More detailed analysis of the product can be achieved using mass spectroscopy. The enzymatic activity of Hct1 may not be confined to a single compound, or conversely, the substrate specificity may be strict, acting on a particular stage in a metabolic

pathway. However, once the substrate is known for Hct1, then a role for the enzyme in the hippocampus may be more directly addressed.

The substrate analysis is dependent on the formal identification of recombinant Hct1 protein produced in yeast. As this can only be done with a specific antibody against Hct1, such antibodies will be raised for Hct1. The antibodies will be of further use in defining Hct1 protein distribution in the hippocampus. Subcellular localisation of Hct1 should contribute to our understanding of Hct1 function. If, for example, Hct1 protein were to be concentrated at the postsynapse, as is the case for CaMKII (Schulman and Lou, 1989), a function in postsynaptic signalling or modification could be addressed. If, in addition, arachidonic acid was found to be a substrate for Hct1, then a role in the production of retrograde messengers could be investigated. The use of an antibody to block Hct1 function or the development of a specific inhibitor of Hct1 activity could be used to assess its involvement in LTP.

Evidence for a role in synaptic plasticity for Hct1 could be addressed further by changes in gene expression. Synaptic activity can be induced in live animals using a number of different paradigms such as seizure or LTP. Changes in gene expression can be detected using *in situ* hybridisation. This has been applied to a number of genes already. The inducible cyclooxygenase, COX-2, (Yamagata *et al.*, 1993) and tissue plasminogen activator (tPA) (Qian *et al.*, 1993) are both immediate early genes which are upregulated after induced seizure, implying a role for these genes in effecting synaptic plasticity. Based on the observation by Stadler *et al.* (1994) that nitric oxide downregulates CYP activity, it might be interesting to determine whether Hct1 expression levels are altered during LTP.

As described in section 6.5, involvement of cytochromes P-450 such as Hct1 in neurodegenerative disease is a possibility worthy of further investigation. Efforts are in progress to clone the human orthologue of Hct1, and to use the cDNA as a probe on human metaphase chromosome preparations. Once the chromosomal location of the

human gene is mapped, the position can be compared with the mapped loci of known human neurodegenerative diseases.

The rat Hct1 promoter may not prove to be useful for transgenic studies, as there is low level expression in brain regions outside the hippocampus. The precise level of this expression is still slightly ambiguous as the signal was close to background levels. Confirmation of the rat expression pattern will however require analysis of reporter gene expression patterns under the control of the isolated rat promoter in transgenic animals. To confirm the mouse Hct1 expression pattern in the adult, during development, and thus assess its suitability as an Hct promoter, an attractive option would be to introduce a reporter gene into the Hct1 locus. A dicistronic targeting vector, such as that described by Mountford *et al.* (1994), can be used to report chromosomal gene expression. This vector inserts a *lacZ* gene prefixed by a picornoviral Internal Ribosome Entry Site (IRES) to permit internal reinitiation of translation. The vector can be introduced into a defined genetic locus using homologous recombination in ES cells, resulting in a mouse expressing β-galactosidase under the transcriptional control of the endogenous gene. This approach is dependent on expression of the gene in ES cells, and Hct1 can be weakly detected in ES cells by Northern analysis (data not shown). The advantages of such a strategy are twofold: the endogenous expression pattern of Hct1 is revealed, while the reporter gene insertion can result in inactivation of the gene product; deficits revealed by crossing such animals to homozygosity would shed light on the biological role of Hct1.

The application of Hct promoters and their practical contribution to the understanding of hippocampal function encourages further characterisation of gene expression in the hippocampus. In addition to more refined versions of differential screens (see Chapter two), an alternative approach to the identification of Hct genes which is proving successful is the analysis of mice with gene trap insertions (for review see Joyner *et al.*, 1992). Gene trap vectors contain a splice acceptor site upstream of a *lacZ*/neomycin fusion gene. Integration into a gene and correct use of the splice acceptor results in production of the fusion

protein in ES cells. Mice generated from these ES cells will bear a marker of expression of the endogenous gene. The analysis of β-galactosidase expression patterns in a number of gene trap lines has uncovered three expression patterns where lacZ staining is highly enriched in hippocampus (personal communication, Dr. M. Steel). One line is of particular interest as expression is largely restricted to the CA areas of the hippocampus with no staining apparent in the dentate gyrus. These mice may be additionally valuable as the *lacZ* insertion not only reveals the expression pattern but may also result in inactivation of the endogenous gene.

There are also recently published reports of genes which may qualify as an Hct promoter. The recent cloning of a serotonin receptor which is expressed in the CA1 of the rat hippocampus and the medial habenulae may be of potential use (Wisden *et al.*, 1993; Erlander *et al.*, 1993). In the intact transgenic animal, effects on the habenulae must be considered, but isolated analysis on transgenic hippocampal slices would be possible. A recent abstract also reported a novel protein tyrosine kinase receptor which is expressed solely in the hippocampus of the adult mouse (Rogers *et al.*, 1994). With further analysis of developmental expression patterns and expression levels, these genes may be of potential use as sources of Hct promoters.

6.8 Conclusion

The development of transgenic technology, and more recently tissue-specific gene inactivation, necessitates the identification of promoters capable of directing defined expression patterns. In the understanding of hippocampus-dependent learning and memory, the use of such promoters to direct modified gene expression or gene inactivation to the hippocampus will contribute to an understanding of the biochemical and molecular mechanisms underlying memory formation.

The identification and preliminary characterisation of Hct1 in the hippocampus has raised a number of questions regarding Hct1 function and its potential role in synaptic plasticity. A possible role in neurodegenerative conditions has also been proposed for Hct1. Indeed,

future experiments may lead to a greater understanding of cytochrome P-450 function in the brain.

CHAPTER SEVEN

MATERIALS AND METHODS

7.1 Materials

Unless otherwise stated, all chemicals were of analytical grade, and supplied by BDH Laboratory Supplies or Sigma. Phenol saturated in Tris buffer was supplied by Fisons Bioscience (P/2318/05). Synthetic oligonucleotides were supplied by Oswel, Scotland. Ethanol was supplied by Hayman. Kodak XAR-5 X-ray film was supplied by H.A. West, Scotland. Radioisotopes were supplied by Amersham and New England Nuclear.

Table 7.1: Standard Solutions

EDTA	0.5 M Na_2EDTA (pH 8.0)
TE	10 mM Tris. Cl (pH 7.4), 1mM EDTA
TAE	1X: 40 mM Tris-acetate, 1 mM EDTA
TBE	1X: 45 mM Tris-borate, 1 mM EDTA
SM (phage storage)	100 mM NaCl, 10 mM $MgSO_4$, 50 mM Tris. Cl (pH 7.5), 0.01% gelatin
SSC	20X: 3 M NaCl, 0.3 M Na citrate
Formamide loading dye	98% formamide, 10 mM EDTA, 0.025% xylene cyanol FF, 0.025% bromophenol blue
Non-denaturing loading dye	4 M urea, 50% sucrose, 50 mM EDTA, 0.025% bromophenol blue

Computer software packages used for DNA and protein analyses were the DNAstar computer package and the University of Wisconsin Genetics Computer Group, (1991), GCG package version 7 (UWGCG).

7.2 Methods

Many of the methods described in this work were based on those described by Sambrook *et al.* (1989).

7.2.1 General Cloning Techniques

(a) Restriction Digestion of DNA

Restriction enzymes and buffers were supplied by Boehringer. The conditions recommended for DNA digestion by the Boehringer catalogue were generally followed.

The products of DNA restriction digests were analysed by electrophoresis using agarose gels (Seakem LE agarose, FMC Bioproducts, 50004) at appropriate concentrations in TAE buffer.

(b) Precipitation of nucleic acids

DNA or RNA was precipitated by the addition of a 0.1 volume 3 M sodium acetate (pH 5.2), followed by 2.5 volumes of -20°C ethanol. The mixture was held at -20°C for 30 minutes, and the nucleic acids pelleted by centrifugation, at 12,000g, 4°C. Very small fragments of DNA (<200 nucleotides) were precipitated as above, with the addition of 1 µl of salmon sperm DNA (20 mg ml^{-1}, Boehringer, 12463420). To remove trace amounts of salt remaining from the precipitation, the nucleic acid pellet was washed in 500 µl of 70% (v/v) ethanol, recovered by centrifugation, air dried, and dissolved in an appropriate volume of dH_2O or TE buffer.

(c) Dephosphorylation of linearised plasmid DNA

Removal of the 5' terminal phosphate groups of linearised plasmid DNA with identical termini was required to prevent recircularisation of the plasmid DNA in subsequent ligation steps.

1. After restriction digestion of DNA, an equal volume of TE was added to the digested DNA.
2. 1 Unit of alkaline phosphatase (Boehringer, 713 023) was added directly, and incubated at 37°C for 15 minutes.
3. EDTA was added to a final concentration of 5 mM, and the phosphatase inactivated by heating the reaction mixture to 75°C for 10 minutes.
4. An equal volume of phenol/chloroform was added, vortexed briefly, and centrifuged at 12,000g before recovering the aqueous phase.
5. The DNA was ethanol precipitated (Section 7.2.1(b)) and resuspend in TE .

(d) Ligation of DNA fragments

Generally 10 nmol of vector DNA was ligated to approximately 50 nmol of insert DNA. A control ligation containing no insert DNA was always included to assess the extent of self-ligation. T4 DNA ligase (Boehringer, 716 359) was added in the presence of 66 mM Tris. Cl (pH 7.5), 5 mM $MgCl_2$, 1 mM DTT, and 1 mM ATP. The reaction components were assembled fresh each time. The reaction was left to proceed overnight at 15°C.

(e) Preparation of competent cells for *E.coli* transformation.

Typically, the efficiency of transformation from such competent cells was 10^6 - 10^7/μg DNA.

1. A fresh overnight culture of *E.coli* cells was diluted 1:100 in LB, and grown at 37°C with shaking until mid-exponential phase (OD_{600} = 0.5).
2. Cells were immediately chilled on ice.
3. Cells were harvested by centrifugation at 3000g for 5 minutes at 4°C.

4. A half volume of ice cold 50 mM $CaCl_2$ was added, and the cells gently resuspended by pipetting. The cells were left on ice for 20 minutes, then harvested (as in step 3).
5. A 1/10 volume of ice cold 50 mM $CaCl_2$ was added, gently resuspending the cell pellet by pipetting. The cells were left on ice for 1-2 hours before use.
6. These cells could be frozen in 15% glycerol and stored at -70°C.

LB medium (per litre): 10 g tryptone (Difco, 0123-01-1)
5 g yeast extract (Difco, 0127-01-7)
5 g NaCl
+/- 15 g agar (Difco, 0140-01)

(f) Transformation of competent *E.coli* cells

1. 100-200 µl of competent cells were aliquoted into a sterile Eppendorf tube for each transformation. (Frozen competent cells were first thawed on ice).
2. DNA was added to each transformation reaction, and left on ice for 40 minutes. Typically, half of the ligation reaction (Section 7.2.1(d)) was used for each transformation.
3. Cells were heat-shocked by placing at 37°C for 2 minutes.
4. The cells were returned to ice and 200 µl of LB medium was added to each. The cells were incubated in a 37°C shaking incubator for 0.5-1 hour to allow the cells to express β-lactamase and confer ampicillin resistance.
5. Cells were spread onto selective medium. Unless specified, most plasmids used conferred ampicillin resistance. Thus 100 µg/ml of ampicillin (Boehringer, 835 269) was added to each LB plate to select for transformants.
6. Transformants appeared after overnight incubation at 37°C.

7.2.2 Nucleic acids isolation

(a) Preparation of plasmid DNA (small scale)

This method was appropriate for a 2-3 ml overnight culture. The resulting DNA could be used for sequencing directly.

1. Cells from a 2-3 ml overnight culture were harvested in 1.5 ml Eppendorf tubes by centrifugation at 12,000g for 2 minutes.
2. All traces of supernatant was removed by aspiration, and the cells resuspended in 100 µl 50 mM glucose, 10 mM EDTA (pH 8.0), 25 mM Tris. Cl (pH 8.0).
3. 200 µl of freshly prepared 0.2 M NaOH, 1% SDS was added, mixed by gentle inversion, and left on ice for 5 minutes.
4. 150 µl of cold 5 M potassium acetate* (pH 5.8) was added, then mixed gently.
5. The mixture was centrifuged in a microfuge at 12,000g for 10 minutes, and the supernatant recovered into a fresh Eppendorf tube.
6. An equal volume of phenol/chloroform was added, vortexed briefly, then centrifuged at 12,000g for 5 minutes.
7. The aqueous phase was recovered into a fresh Eppendorf tube, and the DNA precipitated by adding 800 µl ethanol. After 5 minutes on ice, centrifuge at 12,000g for 5 minutes.
8. The resulting pellet was resuspended in 50 µl of 10 µg/ml RNaseA (Sigma, R-5000) in TE, and incubated at 37°C for 30 minutes.
9. After RNase treatment, the plasmid DNA was purified by adding 30 µl 20% PEG 8000, 2.5 M NaCl, and allowing to precipitate at 4°C overnight (1 hour - overnight).
10. The PEG mixture was centrifuged at 12,000g in a microfuge for 15 minutes, and the pellet washed thoroughly with 2-3 changes of 70% ethanol.
11. The pellet was dried, then resuspended in 20 µl TE.

*5 M potassium acetate was prepared by dissolving 73.5 g of potassium acetate and 28.8 ml of acetic acid in water to a final volume of 250 ml. The resulting solution was 5 M with respect to acetate, and 3 M for potassium.

(b) Preparation of Plasmid DNA (large scale)

Large scale purification of plasmid DNA was prepared from a 250 ml bacterial culture grown in LB*. A Qiagen Plasmid Purification Kit (Qiagen, 12162) was used to purify the plasmid DNA.
Typically, 0.5-1 mg of plasmid DNA was recovered and was quantitated by absorbance at λ = 260 nm, assuming 1 O.D. Unit = 50 µg/ml.

*Terrific Broth, an enriched medium, was substituted for LB when the DNA inserted into the vector was very large and sometimes unstable. (see Sambrook *et al.*, 1989)

(c) Preparation of bacteriophage *lambda* DNA (small scale)

1. To 0.6 ml of a phage liquid lysate, an equal volume of a DEAE-cellulose slurry equilibrated with LB medium was added. (Phage lysate titre must be at least 10^{10} pfu/ml).
2. The slurry and lysate were mixed by gently inverting the tube 20-30 times.
3. The DEAE-cellulose was removed by centrifugation in a microfuge at 12,000g for 5 minutes. The supernatant was transferred to a fresh tube and centrifuged again to remove any remaining traces of DEAE-cellulose.
4. To 0.8 ml of the resulting supernatant was added 100 µl of 5 M NaCl and 540 µl of isopropanol, then chilled at -20°C for 20 minutes.
5. The isopropanol precipitate was centrifuged for 5 minutes at 12,000g in a microfuge.
6. The pellet was washed with 1 ml of 70% ethanol, then centrifuged for 5 minutes at 12,000g to remove all traces of supernatant.
7. The pellet was resuspended in 200 µl TE.
9. An equal volume of phenol was added, and vortexed briefly.
10. The mixture was centrifuged at 12,000g in a microfuge for 5 minutes.
11. The upper aqueous phase was transferred to a fresh Eppendorf tube containing 200 µl of phenol, and vortexed briefly.
12. The mixture was centrifuged at 12,000g in a microfuge for 5 minutes, and the upper aqueous phase recovered.

13. The DNA was ethanol precipitated, washed with 70% ethanol and air-dried (Section 7.2.1(b)).
14. The pellet was resuspended in 20 µl of TE. This yielded enough DNA for one restriction enzyme reaction.

Preparation of DEAE-cellulose (Whatman, 4057 050):
DEAE-cellulose was suspended in several volumes of 0.05 M HCl. With stirring, 10 M NaOH was added to pH 6.8 The resin was left to settle, the supernatant decanted, then resuspended in LB. This was repeated until the pH is 6.8. The slurry was then resuspended in a final ratio of about 75% resin and 25% LB. 0.1% sodium azide was added and stored at 4°C.

(d) Preparation of Bacteriophage *lambda* DNA (large scale)

1. 500 ml of LB medium was inoculated with 1 ml from an overnight culture of an appropriate bacterial strain. The diluted culture was incubated at 37°C with vigorous shaking until the culture had reached an OD_{600} of 0.5.
2. The culture was inoculated with 10^{10} pfu of a phage lysate. The incubation at 37°C with shaking was continued until the culture had lysed. This was typically 3-4 hours.
3. 10 ml of chloroform was added, and the incubation at 37°C with shaking continued for 10 minutes.
4. The lysate was centrifuged at 4000 rpm for 5 minutes (Sorvall GS3) to remove the bacterial particulate matter.
5. DNase (Boehringer, 104 159) and RNaseA (Sigma, R-5000)) were each added to a final concentration of 1 µg/ml, and incubated at room temperature for 30 minutes.
6. Solid NaCl was added to a final concentration of 1 M (29.2g/500ml).
7. Once the NaCl was dissolved, solid PEG 8000 was added to the lysate to a final concentration of 10% (w/v). The PEG was dissolved by swirling.
8. The lysate was left at 4°C overnight.
9. The phage were pelleted by centrifugation at 11,000g for 15 minutes.

10. The phage pellet was resuspended in 5 ml of SM buffer by gentle pipetting, then centrifuged at 7500g for 15 minutes to remove debris.
11. The supernatant was recovered, and layered carefully onto a CsCl step gradient (Figure 7.1).

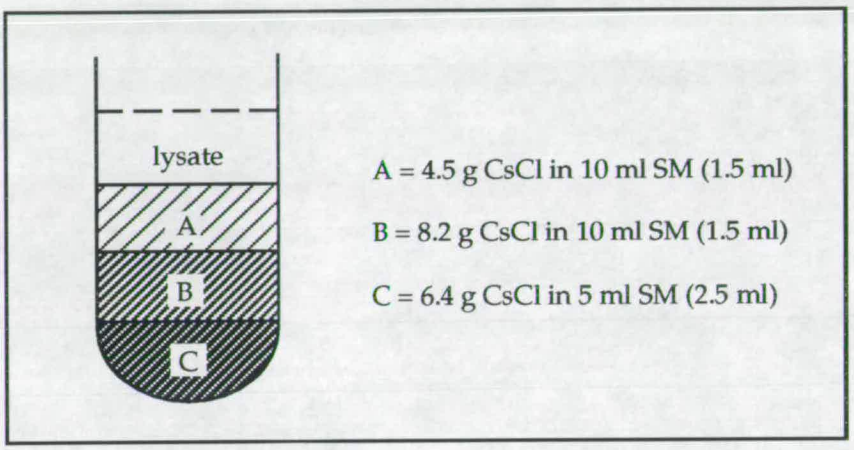


Figure 7.1: CsCl step gradient

12. The CsCl gradient was centrifuged at 11,000g for 90 minutes at 4°C in a Beckman SW41 rotor.
13. The resulting band of bacteriophage particles were recovered by puncturing the side of the tube with a 21 gauge needle. The phage band appeared at the interface of the first two layers as a bluish band.
14. The phage particles were dialysed overnight in 10 mM NaCl, 50 mM Tris. Cl (pH 8.0), 10 mM $MgCl_2$.
15. The dialysed phage were transferred to an Eppendorf tube. An equal volume of phenol was added to the phage and vortexed thoroughly. The mixture was centrifuged at 12,000g in a microfuge for 5 minutes, and the upper aqueous phase transferred to a fresh tube. This phenol extraction was repeated twice.
16. After the third extraction, an equal volume of chloroform was added to the aqueous phase and vortexed thoroughly.
17. The mixture was centrifuged at 12,000g in a microfuge for 5 minutes.

18. The aqueous phase was recovered to a fresh tube.
19. The phage DNA was hooked out of the aqueous phase by layering 2 volumes of absolute ethanol onto the DNA solution, and, using a glass hook, spooling the DNA onto the hook as it mixed with the ethanol.
20. The DNA-covered glass hook was snapped off into an Eppendorf tube containing 1 ml of TE. The DNA was allowed to dissolve by rotating the tube overnight at 4°C.

(e) Preparation of genomic DNA

1. Fresh or frozen tissue was homogenised in guanidinium thiocyanate solution (see section 7.2.2 (g)) using an Ultra-Turrax T25 homogeniser at high speed for 1 minute. Generally, 0.5 g of tissue was homogenised in 10 ml of guanidinium solution. If the homogenate was too viscous, more guanidinium solution was added.
2. The homogenate was gently mixed at room temperature for approximately 1 hour. The homogenate could be stored at 4°C at this point.
3. 1 ml of the homogenate was layered at the bottom of 3 ml of ethanol.
4. Using a glass hook made from a Pasteur pipette, the homogenate was looped up into the ethanol, allowing the genomic DNA to adhere to the glass. This was done repeatedly until the ethanol/guanidinium interface had disappeared.
5. The tip of the hook with the DNA, appearing as white strands, was snapped off into an Eppendorf tube containing 500 µl of TE.
6. The DNA dissolved into the TE by rotating the tube overnight at room temperature. Yields varied between amount and type of tissue. DNA was quantitated by absorbance at λ = 260 nm, assuming 1 O.D. Unit = 50 µg/ml.

(f) General Considerations in the preparation of RNA.

During the preparation and manipulation of RNA, all plasticware was sterile and glassware baked at 180°C. Gloves were used throughout, and all solutions were treated with 0.1% DEPC (Sigma D-

5758) and then autoclaved. Solutions containing Tris.Cl were not treated with DEPC.

(g) Total RNA from mammalian tissues

This method is based on that described in Chirgwin *et al.* (1979).

1. Frozen or fresh tissue samples were homogenised in guanidinium thiocyanate solution, using a Ultra-Turrax T25 homogeniser at high speed for 0.5 - 1 minute. Usually, 0.25 - 0.5 g of tissue was homogenised in 6.ml of guanidinium solution per centrifuge tube.
2. DNA in the homogenate was sheared by passing the homogenate through a 21G needle attached to a syringe several times.
3. 6 ml of 5.7 M CsCl, 10 mM EDTA was aliquoted into each centrifuge tube (Beckman polyallomer tubes, 331372), onto which was carefully layered the homogenate.
4. Samples were centrifuged in a Beckman SW41 rotor for 15 hours, at 30,000 rpm, at 15^{o}C.
5. The supernatants were aspirated, firstly removing the interface, and then all the guanidinium phase. The 'clean' CsCl was then removed until approximately 0.5 ml remained. The tubes were then inverted to drain.
6. Once thoroughly drained, the tube was cut leaving the bottom part containing the RNA pellet.
7. The RNA pellet was washed by gently adding 200 µl of 70% ethanol and then immediately inverting the tube.
8. The RNA was briefly air-dried, then resuspended in 200µl of guanidinium solution. This required a lot of pipetting to resuspend the RNA.
9. The RNA was transferred to an Eppendorf tube containing 400 µl of isopropanol. Any remaining RNA was recovered by a second 200 µl of guanidinium solution, which, too, was added to the isopropanol.
10. The precipitated RNA was washed with 70% ethanol, air-dried briefly, and resuspended in 400 µl of TE.

11. The RNA was added to an Eppendorf tube containing 400 µl of hot phenol (65°C). The mixture was incubated at 65°C for 2 minutes with frequent vigorous vortexing.
12. Samples were centrifuged at 12,000g for 5 minutes. The upper aqueous phase was transferred to a fresh tube containing 400 µl of 65°C phenol, and the extraction repeated.
13. After the second extraction, the aqueous phase was transferred to a fresh tube containing 400 µl of phenol/chloroform at room temperature. The samples were vortexed thoroughly, and centrifuged at 12,000g for 5 minutes.
14. The aqueous phase was transferred to a tube containing 400 µl of chloroform at room temperature. Samples were vortexed thoroughly, and the aqueous phase recovered by centrifugation at 12,000g for 5 minutes.
15. The RNA was ethanol precipitated, washed in 70% ethanol, and air-dried (Section 7.2.1(b)).
16. The RNA pellet was resuspended in TE. Quantitation was by absorbance at λ = 260 nm, assuming 1 O.D. Unit = 40 µg/ml. Yields varied depending on tissue type and amount.

Guanidinium Thiocyanate Solution:

4 M guanidinium thiocyanate (Fluka, 50990)
25 mM sodium citrate
10 mM N-lauryl sarcosine
20 mM β-mercaptoethanol (Sigma stock is 14.2 M)
a few drops of antifoam

(h) mRNA isolation from mammalian tissues.

mRNA from the hippocampus of rats or mice was extracted using the Fast Track mRNA isolation kit Version 3.1 (Invitrogen Cat. K1593-02).

For all other tissues described, the following mRNA isolation method was used.

1. Approximately 0.5 g of fresh or frozen tissue was homogenised for 0.5 - 1.0 minutes at high speed using an Ultra-Turrax T25 homogeniser, in freshly prepared STE/SDS buffer containing 300 µg/ml proteinase K (Sigma, P-0390).
The proteinase K was predigested at 37°C for 1 hour before use.
2. The homogenate was incubated at 37°C for 1-2 hours with occasional swirling.
3. The NaCl concentration of the homogenate was adjusted to 0.5 M, followed by addition of 0.5 - 1.0 ml (packed volume) of pretreated oligo-(dT)-cellulose (Boehringer, 808 229).
The oligo-(dT)-cellulose was pretreated by suspending in 0.1 M NaOH, and rotating for 20 minutes at room temperature. The oligo-(dT)-cellulose was then washed six times in DEPC-treated water.
4. The homogenate was incubated at room temperature for 1 - 2 hours with vigorous shaking.
5. The oligo-(dT)-cellulose suspension was centrifuged at 3000g for 5 minutes, and the supernatant removed.
6. The oligo-(dT)-cellulose was resuspended in 5 ml of prewarmed Binding Buffer by gentle shaking, then allowed to settle.
7. The supernatant was applied to a pretreated Biorad 'Econo-column' (731-1550), equilibrating the column to the high salt condition, and allowed to drain.
The column was pretreated by filling with 0.1 M NaOH, allowing to drain, and then washing through with 6 volumes of DEPC-treated water.
8. The oligo-(dT)-cellulose was resuspended in 1 ml of binding buffer and was applied to the column.
9. A further 3 x 1 ml of binding buffer was applied to the column, draining inbetween each application.
10. 2 x 1 ml of prewarmed wash buffer was applied to the column, and allowed to drain.
11. The RNA was eluted by applying 4 x 1 ml of elution buffer to the column, allowing the column to drain between applications. The eluate was collected into a tube containing 4 ml of phenol/chloroform.
12. The eluate was gently mixed, then centrifuged at 3000g for 5 minutes.

13. The upper aqueous phase was transferred to a tube containing 4 ml of chloroform, gently mixed, and centrifuged at 3000g for 5 minutes.
14. The aqueous phase was transferred to a 15 ml corex tube containing 400 µl of 3 M sodium acetate and 8.8 ml of ethanol.
15. The RNA was precipitated overnight at -70°C, recovered by centrifugation at 10,000 rpm (Sorvall, HB-4) for 30 minutes.
16. The RNA pellet was resuspended in DEPC-treated water, and stored at -70°C. The RNA was quantitated by absorbance at λ = 260 nm, assuming 1 O.D. unit = 40 µg/ml.

STE/SDS:

0.1 M NaCl
10 mM Tris. Cl (pH 7.4)
1 mM EDTA
0.5% SDS

Binding Buffer (prewarmed to 37°C):

0.5 M NaCl
10 mM Tris. Cl (pH 7.4)
1 mM EDTA
0.1% SDS

Wash Buffer (prewarmed to 37°C):

0.1 M NaCl
10 mM Tris. Cl (pH 7.4)
1 mM EDTA
0.1% SDS

Elution Buffer (prewarmed to 65°C):

10 mM Tris. Cl (pH 7.4)
0.1 mM EDTA
0.1% SDS

(i) Total RNA from yeast

This method is based on that of McInerny *et al.*(1994).

1. 25-50 ml cultures of mid-exponential stage of growth yeast cells were prepared and harvested by centrifugation at 3000g for 10 minutes in a 50 ml Falcon tube.
2. Each pellet was resuspended in 1 ml of sterile saline and transferred to an Eppendorf tube; cells were pelleted by centrifugation at 12,000g, and the supernatant discarded (cells could be stored at -70°C at this stage).
3. Cells were resuspended in 75 µl of STE.
4. Hydrochloric acid-washed beads (425-600 micron, Sigma G-9268) were added to just beneath meniscus, and the cells briefly vortexed.
5. 600 µl of NTES was added, and the mixture briefly vortexed again.
6. 500 µl of hot phenol (65°C) was added, and the mixture immediately vortexed.
7. The mixture was incubated at 65°C for 5 minutes with frequent vortexing.
8. The mixture was centrifuged for 1 minute at 12,000g, and the aqueous phase was removed to a second 500 µl aliquot of hot phenol. This was incubated at 65°C for 2 minutes with frequent vortexing.
9. The mixture was centrifuged for 1 minute at 12,000g, and again the aqueous phase was removed to a third 400 µl aliquot of hot phenol.
10. The mixture was incubated at 65°C for two minutes with frequent vortexing.
11. The aqueous phase was removed into 400 µl of phenol/chloroform at room temperature, vortexed, and centrifuged at 12,000g for 1 minute.
12. The aqueous phase was re-extracted with 300 µl of chloroform, vortexed, and centrifuged for 1 minute at 12,000g.
13. Aqueous phase was ethanol precipitated (Section 7.2.1(b)).
14. The RNA pellet was resuspended in 55 µl of ice cold DEPC water, and stored at -70°C.

Sterile saline: 0.9% (w/v) NaCl, autoclaved.

STE: 0.32 M sucrose
20 mM Tris. Cl (pH 7.5)

10 mM EDTA (pH 8.0)
0.5 mg/ml heparin (Sigma H-7005) added just prior to use.

NTES: 100 mM NaCl
5 mM EDTA (pH 8.0)
50 mM Tris. Cl (pH 7.5)
1% (w/v) SDS
0.5 mg/ml heparin added just prior to use.

(j) Southern Blot

This method was based on that described by Southern *et al.* (1975)

1. After electrophoresis, the gel was stained in 0.5 µg/ml ethidium bromide in dH_2O for 30 minutes, after which the gel was destained in dH_2O for the same time.
2. The gel was photographed, and then denatured in 0.5 M NaOH, 1.5 M NaCl for 60 - 90 minutes.
3. Following denaturation, the gel was neutralised in two changes of 1 M Tris. Cl (pH 7.4), 1.5 M NaCl for 60 minutes each.
4. The gel was rinsed briefly in 6XSSC, and then transferred by capillary blot (Figure 7.2) overnight. The membrane used in all transfers was Hybond-N (Amersham, RPN.303N). The transfer buffer was 20XSSC.
5. After the transfer, the filter was rinsed briefly in 2XSSC, then baked at 80°C for 2 hours.

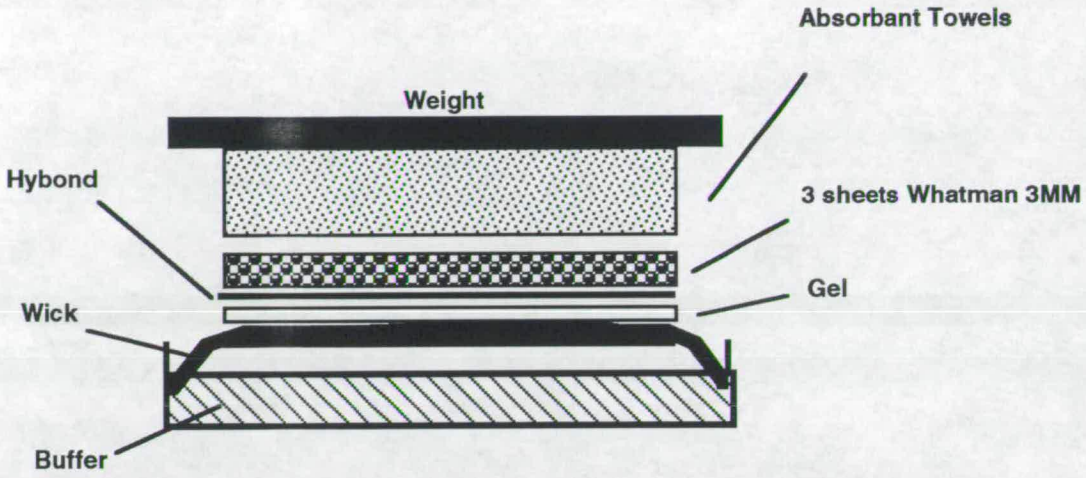


Figure 7.2: Construction of Northern or Southern Blot.

(k) Northern Blot

1. RNA samples were ethanol precipitated (Section 7.2.1(b)).
2. The RNA samples were resuspended in 1xMOPS, 50% (v/v) formamide, 7% (v/v) formaldehyde and 1 µg/ml ethidium bromide, and heated at 65°C for 15 minutes.
3. Formamide loading dye (Table 7.1) was added to each sample before being loaded onto a 1% agarose gel containing 1xMOPS and 7% (v/v) formaldehyde.
4. The gel was electrophoresed for 2-3 hours at 10V/cm, until the bromophenol blue was 12 -14 cm from the origin.
5. After electrophoresis and photography, the gel was ready to blot directly. The capillary transfer of RNA to a Hybond-N (Amersham, RPN.303N) membrane was exactly the same as for DNA (see Figure 7.2).

10 x MOPS 0.4 M MOPS (pH7.0)
100 mM sodium acetate
10 mM EDTA (pH8.0)

7.2.3 Radiolabelling of DNA

(a) Oligonucleotide-labelling of DNA fragments

This method was from Feinberg and Vogelstein (1984).

DNA fragments for radiolabelling were firstly excised from agarose gels. The DNA was purified using the Geneclean II kit (Bio 101, 3106).

1. 10-25 ng of DNA in a final volume of 31 µl of water was denatured by heating at 95°C for 2 minutes.
2. Tubes were placed on ice, then centrifuged briefly in a microfuge to collect the contents to the bottom.
3. The following 50 µl reaction was assembled:

DNA	31 µl
"OLB"	10 µl
10mg/ml BSA	2 µl (Boehringer, 711 454)
α-^{32}P-dCTP	50 µCi (3000Ci/mmol)
Klenow	2 Units (Boehringer, 1008 412)

4. The reaction was incubated at room temperature for 1-2 hours (up to overnight).
5. The reaction was stopped by adding 0.1 M EDTA.
6. Labelled DNA was purified from unincorporated α-^{32}P-dCTP using the spun column method.

"OLB":

Solution A: 625 µl 2M Tris.Cl (pH8.0), 25 µl 5M $MgCl_2$, 350 µl dH_2O, 18 µl β-mercaptoethanol, 5 µl each of 0.1 M dATP, dTTP and dGTP

Solution B: 2 M Hepes (pH 6.0). Store at 4°C

Solution C: Hexadeoxyribonucleotides (Pharmacia, 27-2166-01) evenly suspended in 3 mM Tris. Cl, 0.2 mM EDTA (pH 7.0) at 90 O.D. units/ml. Store at -20°C

"OLB" is a mixture of solutions A, B and C in a ratio of 2:5:3.

(b) T4 polynucleotide kinase-labelling of 5' termini.

Oligonucleotides could be labelled directly using this method, however DNA fragments would first be dephosphorylated.

10 pmoles of oligonucleotide was combined with 50 mM Tris.Cl (pH8.2), 10 mM $MgCl_2$, 0.1 mM EDTA, 5 mM DTT, 0.1 mM spermidine (10X buffer supplied by Boehringer), and 50 µCi of γ-^{32}P-ATP (5000 Ci/mmol). 10 Units of T4 polynucleotide kinase (Boehringer, 174 645) was added and incubated at 37°C for 45 minutes. The kinase was subsequently inactivated by heating at 68°C for 10 minutes. The labelled oligonucleotide was purified by ethanol precipitation with 10 µg of yeast tRNA (Section 7.2.1(b)), and resuspended in 100 µl of dH_2O.

(c) Terminal Transferase tailing of 3' termini.

This method was used to label oligonucleotides with a tail of ^{35}S for *in situ* hybridisation.

1. 10 ng of oligonucleotide was tailed in a 20 µl reaction. The following were added in order:
   - 3 µl dH_2O
   - 2 µl oligonucleotide (0.5 - 1.0 nmol)
   - 4 µl 5X terminal transferase buffer (supplied by Boehringer)
   - 6 µl 5 mM $CoCl_2$
   - 25 Units terminal transferase (Boehringer, 220 582)
   - 40 µCi α-^{35}S-dATP (1000 Ci/mmol)
2. The reaction was incubated at 37°C for 1 hour.
3. 80 µl of TE and 20 µg yeast tRNA were added to quench the reaction.
4. 100 µl of phenol/chloroform was added, vortexed briefly, then centrifuged at 12,000 g in a microfuge for 5 minutes.
5. The upper aqueous phase was transferred to a screw-capped microfuge tube containing 10 µl 3 M sodium acetate and 250 µl of absolute ethanol.
6. The labelled oligonucleotide was precipitated, washed in 70% ethanol and dried (Section 7.2.1(b)).
7. The labelled oligonucleotide was resuspended in 50 µl TE*.

* If the labelled oligonucleotide was to be used for *in situ* hybridisation, the TE for resuspension contained 10 mM DTT.

7.2.4 Hybridisation Conditions.

Hybridisation conditions used were based on those described by Church and Gilbert (1984).

1. Filters were pre-wet in 2XSSC.
2. The hybridisation was performed in a rotating glass cylinder (Techne Hybridiser ovens). 10 ml of Hybridisation Buffer was added to the cylinder with the filter.
3. Prehybridisation and hybridisation were carried out at 68°C unless otherwise specified.
4. The filters were prehybridised for 30 minutes, after which the probe was added directly and hybridisation proceeded overnight. (Double-stranded probes were denatured by boiling for 2 minutes, then placing on ice).
5. Washes were performed at 68°C (unless otherwise stated) with 2 changes of Wash Buffer I for 10 minutes each, followed by three changes of Wash Buffer II each for 20 minutes.
6. The filters were blotted dry, but not allowed to dry out, then placed between Saran wrap, and against X-ray film for autoradiography.

Hybridisation Buffer:

0.25 M sodium phosphate pH 7.2
1 mM EDTA
7% SDS
1% BSA

Wash Buffer I:

20 mM sodium phosphate pH 7.2
2.5% SDS
0.25% BSA

1 mM EDTA

Wash Buffer II:

20 mM sodium phosphate pH 7.2
1 mM EDTA
1% SDS

7.2.5 Screening of Bacteriophage *lambda* libraries.

The rat hippocampus cDNA library was oligo-(dT)-NotI primed and cloned in *lambda* ZAP II (Stratagene) with an EcoRI adaptor at the 5' end, and was prepared in the lab by Miss M. Richardson and Dr. J. Mason.

The mouse liver cDNA library was oligo-(dT)-primed and cloned into *lambda* gt10 with EcoRI/NotI adaptors, and was a gift from Dr. B. Luckow, Heidelberg.

The mouse ES cell genomic library was cloned from a partial Sau3A digest into *lambda* DASH II (Stratagene), and was a gift from A. Reaume, Toronto.

1. A 1:50 diluted culture of an appropriate bacterial strain was grown to mid-log phase (OD_{600} = 0.5) in LB supplemented with 0.2% maltose and 10 mM $MgSO_4$.
2. The cells were chilled on ice, harvested by centrifugation in a screw-capped Falcon tube at 3000g for 5 minutes, then resuspended in 0.5 volume of ice-cold 10 mM $MgSO_4$.
3. 600 µl of cells were infected with 50,000 recombinant phage (diluted in SM), incubating at 37°C for 15 minutes.
4. 10 ml of 0.7% agarose in LB, supplemented with 10 mM $MgSO_4$ (melted and warmed at 45°C) was added. The tube was inverted slowly to mix, and then poured evenly over a dried LB/agar plate (15 cm diameter).

5. Once the top layer was set, the plates were inverted and incubated at 37°C until the appearance of plaques in the bacterial lawn. This time varied between 7 - 10 hours.
6. The plaques were grown until they were touching each other but not confluent. The plates were then chilled.
7. Replica lifts of the phage were made by layering a piece of Hybond-N nylon membrane (Amersham RPN.132N) onto the plaques for 1 minute. The orientation of the filter was marked. The second lift was layered on to the plaques for 2 minutes, and the same orientation marked.
8. The filter was lifted off carefully with forceps and placed in 0.5 M NaOH, 1.5 M NaCl for 4 minutes.
9. The filters were neutralised in 1 M Tris. Cl (pH 7.4), 1.5 M NaCl for 4 minutes.
10. The filters were washed briefly in 2XSSC and then air-dried, before baking for 2 hours at 80°C.
11. The filters were then hybridised with the selected probe (Section 7.2.4).

In vivo excision of pBluescript from *lambda* ZAP II vector was performed using the ExAssist/SOLR system (Stratagene, 200253).

7.2.6 Dideoxy-termination sequencing of double-stranded DNA templates.

1. Approximately 2 µg of uncut plasmid DNA was denatured by incubating in 200 mM NaOH for 5 minutes at room temperature. The DNA was precipitated with 1/5 volume of 10 M ammonium acetate and 4 volumes of absolute ethanol at -20°C for 15 minutes.
2. The DNA precipitate was washed with 1 ml of 70% ethanol, dried and resuspended in 7 µl of water.
3. The sequencing reactions were prepared using the Sequenase Version 2.0 kit (USB, 70770).
4. Products were electrophoresed through a 0.2 mm 6% denaturing (i.e. 8M urea) acrylamide gel in TBE.

5. After electrophoresis, the gel was soaked in 10% methanol, 10% acetic acid for 10 minutes to leach out the urea and to fix the DNA. The gel was dried at 80°C under vacuum.
6. Sequencing products were visualised by autoradiography using Kodak XAR-5 film.

To sequence the full length of a DNA insert, nested deletions in the inserted DNA were made using the Erase-a-Base System (Promega, E5750), with the following modifications:
(i) Before the ligation to circularise the deletion-containing vectors, the deleted DNA was purified by agarose gel electrophoresis, and the excised DNA purified using GeneClean. This was found to increase the proportion of the correct size deletion.
(ii) In each case, Bluescript II (KS and SK) were used. The restriction sites in the polylinker used for Exonuclease III-sensitive and -insensitive cleavage were SalI and KpnI respectively.
(iii) Competent XL1 blue cells (Section 7.2.1(e)) were used for the transformation.

7.2.7 *In situ* hybridisation

Frozen 10 μ coronal sections of rat and mouse brains were provided by Dr. M. Steel.

(a) Preparation of frozen sections for *in situ* hybridisation

All manipulations were carried out at room temperature. All solutions were 0.1% (v/v) DEPC-treated and autoclaved. Microscope slides were placed back-to-back in Coplin jars, with the following solutions poured on or off for each step. For each step, 30-35 ml of each solution per Coplin jar was required.

1. Sections were fixed in 4% (w/v) paraformaldehyde in PBS for 10 minutes.
2. Sections were rinsed in two changes of PBS for 5 minutes.

3. Sections were proteinase K treated by immersing in 20 µg/ml proteinase K in 50 mM Tris. Cl (7.4), 5 mM EDTA for 5 minutes.
4. Sections were rinsed in PBS for 5 minutes.
5. Sections were fixed again in 4% (w/v) paraformaldehyde in PBS for 10 minutes. (The same solution was reused from step 1).
6. Sections were rinsed in water for 5 minutes.
7. The sections were acetylated by placing in 0.25% (v/v) acetic anhydride* for 10 minutes.
8. The sections were rinsed briefly in PBS.
9. The sections were rinsed in saline for 5 minutes.
10. The sections were dehydrated by passing through each of the following ethanol concentrations:

30% ethanol, for 1 minute
50% ethanol, for 1 minute
70% ethanol, for 5 minutes
85% ethanol, for 1 minute
95% ethanol, for 1 minute
100% ethanol, for 1 minute
100% ethanol, for 1 minute.

11. Sections were immersed in chloroform for 5 minutes, rinsed briefly in two changes of absolute ethanol, and a final rinse in 95% ethanol.
12. The slides were removed from the Coplin jars and allowed to air dry before hybridisation.

Paraformaldehyde: PBS was heated to 65°C on a stirrer, then the paraformaldehyde was added to 4% (w/v). 5 M NaOH was added to help the paraformaldehyde to dissolve (approximately 5 drops). Once in solution, the solution was cooled to room temperature and adjusted to pH 7.4 with concentrated HCl.

Saline: 0.8% (w/v) NaCl

Acetic anhydride: 0.25% (v/v) was added to 0.1 M triethanolamine in 0.8% (w/v) NaCl.

*(Acetic Anhydride is very unstable and was made as close to usage time as possible).

(b) Hybridisation Conditions

All probes were oligonucleotides which were labelled by homopolymer tailing using α-^{35}S-dATP and terminal transferase as described (Section 7.2.3 (c))

The sequences or references of the oligonucleotides used as probes for *in situ* hybridisation were as follows:

rat Hct1 (a 45-mer, beginning 26 nt 5' from the polyA tail, nucleotides 1361-1403 in Figure 4.2) (for relative position in mouse gene, see Figure 4.3)
5'-GACAGGTTTTGTGACCCAAAACAAACTGGATGGATCGCAA TC-3'

mouse Hct1 (nt 1558-1599) (also see Figure 4.3)
5'-ATCACGGAGCTCAGCACATGCAGCCTTACTCTGCAAAGCTTC-3'

rat clone 13 (a 42-mer, beginning 112 nt 5' from polyA tail)
5'-TATATCCATACCAACTTATTGGGAGTCCCATCCTACCTCATCA GC-3'

rat/mouse muscarinic receptor M1 (Buckley *et al.*, 1988)

rat/mouse opsins (Nathans *et al.*, 1986)

1. The prepared ^{35}S -tailed probe (resuspended in 10 mM DTT in TE) was diluted to 2 x 10^6 cpm/ml in hybridisation buffer. DTT is also added to this mixture to a final concentration of 50 mM.
2. 100 µl of the probe mixture was carefully layered onto each microscope slide. A piece of parafilm cut to the size of the microscope slide was then layered over the probe mixture, allowing the probe and

hybridisation mixture to cover all the sections. Air bubbles under the parafilm were avoided.

3. The slides were placed in a humidified container, sealed, and incubated at 37°C overnight.

4. After hybridisation, the parafilm was carefully removed using forceps.

5. The slides were placed back in Coplin jars, and the hybridised sections washed in four changes of 1XSSC for 15 minutes at 55°C or 60°C, and then two changes of 1XSSC for 30 minutes at room temperature.

6. The slides were rinsed briefly in dH_2O, then left to air dry.

Hybridisation Buffer*:

4XSSC
50% (v/v) deionised formamide
10% (w/v) dextran sulphate
1X Denhardt's solution
0.1% (w/v) SDS
500 µg/ml ssDNA
250 µg/ml yeast tRNA
*buffer was de-gassed before use

(c) Autoradiography

Detection of hybridisation signals was by autoradiography. Slides were placed against Kodak XAR-5 film for 2-7 days for a low resolution image. For a high resolution image, a photographic liquid emulsion (LM-1, Amersham RPN.40) was used. All manipulations were carried out under safelight conditions.

1. The emulsion was melted by placing in a 43°C water bath.

2. Once melted, the emulsion was poured into a dipping vessel.

3. Blank microscope slides were first dipped into the emulsion until no bubbles could be seen on the slides.

4. Each test slide was dipped evenly into the emulsion vertically for 1-2 seconds.

5. Slides were withdrawn from the emulsion, allowed to drain briefly then the back of each wiped with a tissue to remove excess emulsion.
6. The slides were laid flat to dry.
7. Once dry, the slides were laid flat in a container containing silica gel, and left at room temperature overnight.
8. The slides were transferred into a plastic slide box containing more silica gel, and sealed with tape. The box was covered with two layers of aluminium foil and stored at 4°C for exposure.
9. When ready for developing, the emulsion-dipped slides were brought back to room temperature.
10. Hybridised sections were placed in D19 developer (Kodak, 502 7065) for 2 minutes, then transferred to 1% (v/v) acetic acid, 1% (v/v) glycerol for 1 minute, then to 30% (w/v) sodium thiosulphate for 2 minutes.
11. The sections were finally washed in a large volume of running water for 10 minutes.

(d) Counterstaining and Mounting of Sections.

1. While the sections were still wet, the slides were placed in 1.0% (w/v) methyl green stain (Sigma, M-8884), and stained for 1-2 minutes.
2. The sections were briefly rinsed in water to remove excess stain and left to air dry.
3. Sections were mounted in DePeX (BDH, 36125), and covered with a glass coverslip.

7.2.8 Transformation of Yeast

This method was from Dr J. Tsang, Dundee.

1. A 100 ml culture (2-6 x 10^6/ml) of budding yeast cells grown in YEPD was prepared. The cells were harvested by centrifugation at 3000g for 5 minutes and washed in TE.
2. The cells were resuspended in 0.1 M lithium acetate in TE to a concentration of approximately 5 x 10^8/ml, and left to shake in a 30°C incubator for 1 hour (1-24 hours).

3. To 100 µl of competent cells, 40-80 µg carrier DNA, and 0.1-5 µg of plasmid DNA was added in a 10 µl volume.
4. Cells were incubated at 30°C for 30 minutes.
5. 0.7 ml of 40% PEG 4000 in 0.1 M lithium acetate was added, and the cells incubated for 30 minutes (0.5-2.5 hours).
6. Cells were heat-shocked at 42°C for 15 minutes.
7. The cells were collected by centrifugation at 5000g in a microfuge for 45 seconds, and then washed twice with TE.
8. The cells were resuspended in 100-200 µl TE, and plated onto selective medium.
9. Transformants appeared after 3-5 days incubation at 30°C.

YEPD: 2% glucose
1% peptone
1% yeast extract

7.2.9 Preparation of yeast total protein extract

This method was from McInerny *et al.* (1994).

1. 25-50 ml cultures of yeast cells in mid-late exponential stage of growth were harvested by centrifugation at 3000g for 5 minutes.
2. The cell pellet could be frozen at -70°C. Cell pellets were resuspended in 100 µl of lysis buffer in Eppendorf tubes. Acid washed glass beads (425-600 micron, Sigma G-9268) were added to just below the meniscus.
3. Cells were broken by vigorous vortexing: 15 minutes of 30 seconds, with 30 second intervals on ice.
4. Cell walls were pelleted by 5 minute centrifugation at 12,000g.
5. Supernatants were transferred to another Eppendorf tube, and the protein extract clarified by centrifugation at 12,000g for 30 minutes at 4°C.
6. Supernatants were transferred to a fresh Eppendorf tube. Samples were snap-frozen on solid CO_2 and stored at -70°C.

Lysis buffer:
50 mM KCl

50 mM Tris. Cl (pH 7.4)
25% (v/v) glycerol
2 mM DTT
0.1% (v/v) Triton X-100
25 µg/ml chymostatin (Sigma, C-7268)
25 µg/ml antipain (Sigma, A-6271)
25 µg/ml leupeptin (Sigma, L-2884)
25 µg/ml aprotonin (Sigma, A-6279)
25 µg/ml pepstatin (Sigma, P-4265)
0.2 mM pefabloc (Boehringer, 1429 868)

REFERENCES

Abdelgadir, S.E., Resko, J.A., Ojeda, S.R., Lephart, E.D., McPhaul, M.J., and Roselli, C.E. (1994) Androgens regulate aromatase cytochrome P-450 messenger ribonucleic acid in rat brain. Endocrinology 135: 395-401

Abeliovich, A., Chen, C., Goda, Y., Silva, A.J., Stevens, C.F., and Tonegawa, S. (1993a) Modified hippocampal long-term potentiation in PKCγ mutant mice. Cell 75: 1253-1262

Abeliovich, A., Paylor, R., Chen, C., Kim, J.J., Wehner, J.M., and Tonegawa, S. (1993b) PKCγ mutant mice exhibit moderate deficits in contextual learning. Cell 75: 1263-1271

Adams, S.M., Sharp, M.G.F., Walker, R.A., Brammar, W.J., and Varley, J.M. (1992) Differential expression of translation-associated genes in benign and malignant human breast tumours. Br. J. Cancer 65: 65-71

Amaral, D.G. (1993) Emerging principles of intrinsic hippocampal organisation. Curr. Opinion. Neurobiol. 3: 225-229

Anandatheerthavarada, H.K., Williams, J.F., and Wecker, L. (1993) The chronic administration of nicotine induces cytochrome P-450 in rat brain. J. Neurochem. 60: 1941-1944

Artola, A., and Singer, W. (1987) Long-term potentiation and NMDA receptors in rat visual cortex. Nature 330: 649-652

Ayer-Lelievre, C., Olson, L., Ebendal, T., Seiger, A., and Persson, H. (1988) Expression of the beta-nerve growth factor gene in hippocampal neurons. Science 240: 1339-1341

Bading, H., and Greenberg, M.E. (1991) Stimulation of protein tyrosine phosphorylation by NMDA receptor activation. Science 253: 912-914

Bashir, Z.I., Bortolotto, Z.A., Davies, C.H., Berretta, N., Irving, A.J., Seal, A.J., Henley, J.M., Jane, D.E., Watkins, J.C., and Collingridge, G.L. (1993) Induction of LTP in the hippocampus needs synaptic activation of glutamate metabotropic receptors. Nature 363: 347-350

Bekkers, J.M., and Stevens, C.F. (1990) Presynaptic mechanism for long-term potentiation in the hippocampus. Nature 346: 724-729

Bernstein, A., and Breitman, M. (1989) Genetic ablation in transgenic mice. Mol. Biol. Med. 6: 523-530

Bhamre, S., Anandatheerathavarada, H.K., Shankar, S.K., Boyd, M.R., and Ravindranath, V. (1993) Purification of multiple forms of cytochrome P450 from a human brain and reconstitution of catalytic activities. Arch. Biochem. Biophys. 301: 251-255

Bjorkhem, I. (1985) Mechanisms of bile acid biosynthesis in mammalian liver. In Sterols and Bile Acids, H. Danielsson and J. Sjovall eds. (Elsevier Science Publishers B.V. Amsterdam) pp. 231-278

Black, S.M., Ellard, S., Meehan, R.R., Parry, J.M., Adesnik, M., Beggs, J.D., and Wolf, C.R. (1989) The expression of cytochrome P450IIB1 in *Saccharomyces cerevisiae* results in an increased mutation frequency when exposed to cyclophosphamide. Carcinogenesis 10: 2139-2143

Bliss, T.V.P., and Lømo, T. (1973) Long lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. J. Physiol. 232: 331-357

Bliss, T.V.P., Errington, M.L., Lynch, M.A., and Williams, J.H. (1990) Presynaptic mechanisms in hippocampal long-term potentiation. Cold Spring Harbor Symp. Quant. Biol. 55: 119-129

Bliss, T.V.P., and Collingridge, G.L. (1993) A synaptic model of memory: long-term potentiation in the hippocampus Nature 361: 31-39

Bornheim, L.M., Kim, K.Y., Chen, B., and Correia, M.A. (1993) The effect of cannabidiol on mouse hepatic microsomal cytochrome P-450-dependent anandamide metabolism. Biochem. Biophys. Res. Comm. 197: 740-746

Bredt, D.S., Glatt, C.E., Hwang, P.M., Fotuhi, M., Dawson, T.M., and Snyder, S.H. (1991a) Nitric oxide synthase protein and mRNA are discretely localized in neuronal populations of the mammalian CNS together with NADPH diaphorase. Neuron 7: 615-624

Bredt, D.S., Hwang, P.M., Glatt, C.E., Lowenstein, C., Reed, R.R., and Snyder, S.H. (1991b) Cloned and expressed nitric oxide synthase structurally resembles cytochrome P-450 reductase. Nature 351: 714-718

Breslow, J.L. (1993) Genetics of lipoprotein disorders. Circulation 87: 1116-1121

Buckley, N.J., Bonner, T.I., and Brann, M.R. (1988) Localization of a family of muscarinic receptor mRNAs in rat brain. J. Neuroscience 8: 4646-52

Byrne, J.H., Baxter, D.A., Buonomano, D.V., and Raymond, J.L. (1990) Neuronal and network determinants of simple and higher-order features of associative learning: experimental and modeling approaches. Cold Spring Harbor Symp. Quant. Biol. 55: 175-186

Byrne, J.H., Baxter, D.A., Buonomano, D.V., Cleary, L.J., Eskin, A., Goldsmith, J.R., McClendon, E.V., Nazif, F.A., Noel, F., and Scholz, K.P. (1991) Neural and molecular bases of nonassociative and associative learning in *Aplysia*. In Activity-Driven CNS Changes in Learning and Development. Volume 627, Annals of the New York Academy of Sciences

Cali, J.J., and Russell, D.W. (1991) Characterisation of human sterol 27-hydroxylase. J. Biol. Chem. 266: 7774-7778

Chen, L., and Huang, L.-Y. M. (1992) Protein kinase C reduces Mg^{2+} block of NMDA-receptor channels as a mechanism of modulation. Nature 356: 521-523

Ching, M.S., Lennard, M.S., Tucker, G.T., Woods, H.F., Kelly, D.E., and Kelly, S.T. (1991) The expression of human cytochrome P450IA1 in the yeast *Saccharomyces cerevisiae*. Biochem. Pharmacol. 42: 753-758

Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J., and Rutter, W.J. (1979) Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry 18: 5294-5299

Cholerton, S., Daly, A.K., and Idle, J.R. (1992) The role of individual human cytochromes P-450 in drug metabolism and clinical response. Trends Pharm. 13: 434-439

Church, G.M., and Gilbert, W. (1984) Genomic sequencing. Proc. Natl. Acad. Sci. USA 81: 1991-1995

Cohen, J.C., Cali, J.J., Jelinek, D.F., Mehrabian, M., Sparkes, R.S., Lusis, A.J., Russell, D.W., and Hobbs, H.H. (1992) Cloning of the human cholesterol 7 alpha-hydroxylase gene (CYP7) and localisation to chromosome 8q11-q12. Genomics 14: 153-161

Cohen, N.J., and Eichenbaum, H. (1993) Memory, amnesia and the hippocampal system. MIT Press

Cole, A.J., Saffen, D.W., Baraban, J.M., and Worley, P.F. (1989) Rapid increase of an immediate early messenger RNA in hippocampal neurons by synaptic NMDA receptor activation. Nature 340: 474-476

Collingridge, G.L., Kehl, S.J., and McLennan, H. (1983) Excitatory amino acids in synaptic transmission in the Schaffer collateral-commissural pathway of the rat hippocampus. J. Physiol. 334: 33-46

Cook, D.G., and Carew, T.J. (1986) Operant conditioning of head waving in *Aplysia*. Proc. Natl. Acad. Sci. USA 83: 1120-1124

Corder, E.H., Saunders, A.M., and Strittmatter, W.J. (1993) Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. Science 261: 921-923

Cote, F., Collard, J-.F., and Julien, J-.P. (1993) Progressive neuronopathy in transgenic mice expressing the human neurofilament heavy gene: a mouse model of amyotrophic lateral sclerosis. Cell 73: 35-46

Davis, R.L. (1993) Mushroom bodies and *Drosophila* learning. Neuron 11: 1-14

Davis, R.L., Weintraub, H., and Lassar, A.B. (1987) Expression of a single transfected cDNA converts fibroblasts to myoblasts. Cell 51: 987-1000

Denny, J.B., Polan-Curtain, J., Ghuman, A., Wayner, M.J., and Armstrong, D.L. (1990) Calpain inhibitors block long-term potentiation. Brain Res. 534: 317-320

de Sauvage, F., Kruys, V., Marinx, O., Huez, G., and Octave, J.N. (1992) Alternative polyadenylation of the amyloid protein precursor mRNA regulates translation. EMBO J. 11: 3099-3103

Devane, W.A., Hanus, L., Breuer, A., Pertwee, R.G., Stevenson, L.A., Griffin, G., Gibson, D., Mandelbaum, A., Etinger, A., and Mechoulam, R. (1992) Isolation and structure of a brain constituent that binds to the cannabinoid receptor. Science 258: 1946-1949

Donahue, B.S., Skottner-Lundin, A., and Morgan, E.T. (1991) Growth hormone-dependent and -independent regulation of cytochrome P-450 isozyme expression in streptozotocin-diabetic rats. Endocrinology 128: 2065-2076

Downward, J. (1990) The ras superfamily of GTP-binding proteins. Trends Biochem. Sci. 15: 469-472

Dumuis, A., Sebben, M., Haynes, L., Pin, J.-P., and Bockaert, J. (1988) NMDA receptors activate the arachidonic acid cascade system in striatal neurons. Nature 336: 68-70

Erlander, M.G., Lovenberg, T.W., Baron, B.M., De Lecea, L., Danielson, P.E., Racke, M., Slone, A.L., Siegel, B.W., Foye, P.E., Cannon, K., Burns, J.E., and Sutcliffe, J.G. (1993) Two members of a distinct subfamily of 5-hydroxytryptamine receptors differentially expressed in rat brain. Proc. Natl. Acad. Sci. USA 90: 1-6

Ernst, M. Gearing, D.P., and Dunn, A.R. (1994) Functional and biochemical association of Hck with the LIF/Il-6 receptor signal transducing subunit gp130 in embryonic stem cells. EMBO J. 13: 1574-1584

Evans, M.J., and Kaufman, M.H. (1981) Establishment in culture of pluripotential cells from mouse embryos. Nature 292: 154-156

Feinberg, A.P., and Vogelstein, B. (1984) A technique for radiolabelling DNA restriction fragments to high specific activity. Anal. Biochem. 137: 266-267

Frey, U., Huang, Y.-Y., and Kandel, E.R. (1993) Effects of cAMP simulate a late stage of LTP in hippocampal CA1 neurons. Science 260: 1661-1664

Froehner, S.C. (1991) The sub-membrane machinery for nicotinic acetylcholine receptor clustering. J. Cell Biol. 114: 1-7

Gadaleta, G., Pepe, G., DeCandia, G., Quagliariello, C., Sbisa, E., and Saccone, C. (1989) The complete nucleotide sequence of the *Rattus norvegicus* mitochondrial genome: cryptic signals revealed by comparative analysis between vertebrates. J. Mol. Evol. 28: 497-516

Gehring, W.J. (1987) Homeo boxes in the study of development. Science 236: 1245-1252

Geng, J., and Strobel, H.W. (1993) Identification of cytochromes P-450 1A2, 2A1, 2C7, 2E1 in rat glioma C6 cell line by RT-PCR and specific restriction enzyme digestion. Biochem. Biophys. Res Comm. 197: 1179-1184

Gonzalez, F. (1992) Human cytochrome P-450: problems and prospects. Trends Pharm. 13: 346-352

Grant, S.G.N., O'Dell, T.J., Karl, K.A., Stein, P.L., Soriano, P., and Kandel, E.R. (1992) Impaired long-term potentiation, spatial learning, and hippocampal development in *fyn* mutant mice. Science 258: 1903-1910

Gu, H., Marth, J.D., Orban, P.C., Mossmann, H., Rajewsky, K. (1994) Deletion of a DNA polymerase β gene segment in T cells using cell-type specific gene targeting. Science 265: 103-106

Gulyas, A.I., Toth, K., Danos, P., and Freund, T.F. (1991) Subpopulations of GABAergic neurons containing parvalbumin, calbindin D28k, and cholecystokinin in the rat hippocampus. J. Comp. Neurol. 312: 371-378

Haglund, L., Kohler, C., Haaparanta, T., Goldstein, M., and Gustafsson, J. (1984) Presence of NADPH-cytochrome P450 reductase in central catecholaminergic neurones. Nature 307: 259-262

Hall, A. (1993) Ras-related proteins. Curr. Opin. Cell Biol. 5: 265-268

Han, J.-W., McCormick, F., and Macara, I.G. (1991) Regulation of Ras-GAP and the neurofibromatosis-1 gene product by eicosanoids. Science 252: 576-578

Haniu, M., Armes, L.G., Yasunobu, K.T., Shastry, B.A., and Gunsalus, I.C. (1982) Amino acid sequence of the *Pseudomonas putida* cytochrome P-450. II. Cyanogen bromide peptides, acid cleavage peptides, and the complete sequence. J. Biol. Chem. 257: 12664-12671

Hansson, T., Tindberg, N., Ingelman-Sundberg, M., and Kohler, C. (1990) Regional distribution of ethanol-inducible cytochrome P450 IIE1 in the rat central nervous system. Neuroscience 34: 451-463

Hardwick, J.P., Song, B-J., Huberman, E., and Gonzalez, F.J. (1987) Isolation, complementary DNA sequence, and regulation of rat hepatic lauric acid ω-hydroxylase (Cytochrome P-450$_{LA\omega}$). J. Biol. Chem. 262: 801-810

Hasty, P., Ramirez-Solis, R., Krumlauf, R., and Bradley, A. (1991) Introduction of a subtle mutation into the Hox-2.6 locus in embryonic stem cells. Nature 350: 243-246

Hilton, D.J. (1992) LIF: lots of interesting functions. Trends Biochem. Sci. 17: 72-76

Hille, B. (1992) G protein-coupled mechanisms and nervous signaling. Neuron 9: 187-195

Hodgson, A.V., White, T.B., White, J.W., and Strobel, H.W. (1993) Expression analysis of the mixed function oxidase system in rat brain by the polymerase chain reaction. Mol. Cell. Biochem. 120: 171-179

Hofer, M., Pagliusi, S.R., Hohn, A., Leibrock, J., and Barde, Y.-A. (1990) Regional distribution of brain-derived neurotrophic factor mRNA in the adult mouse brain. EMBO J. 8: 2459-2464

Hooper, M.L. (1992) Embryonal stem cells: introducing planned changes into the animal germline. (Harwood Acad. Publishers)

Hyman, B.T., Van Hoesen, G.W., Damasio, A.R., and Barnes, C.L. (1984) Alzheimer's disease: Cell-specific pathology isolates the hippocampal formation. Science 225: 1168-1170

Imai, Y., and Nakamura, M. (1989) Point mutations at threonine-301 modify substrate specificity of rabbit liver microsomal cytochromes P-450 (laurate (ω-1)-hydroxylase and testosterone 16α-hydroxylase). Biochem. Biophys. Res. Comm. 158: 717-722

Joyner, A.L., Auerbach, A., and Skarnes, W.C. (1992) The gene trap approach in embryonic stem cells: the potential for genetic screens in mice. In Postimplantation development in the mouse. Wiley, Chichester (Ciba Foundation Symposium 165) pp227-297

Kandel, E.R., Schwartz, J.H., and Jessell, T.M. (eds) (1991) Principles of neural science. Third Edition Elsevier

Kemper, B., and Szczesna-Skorupa, E. (1989) Cytochrome P-450 membrane signals. Drug Metab. Rev. 20: 811-820

Kirchhausen, T., Nathanson, K.L., Matsui, W., Vaisberg, A., Chow, E.P., Burne, C., Keen, J.H., and Davis, A.E. (1989) Structural and functional division into two domains of the large (100- to 115-kDa) chains of the clathrin-associated protein complex AP-2. Proc. Natl. Acad. Sci USA 86: 2612-2616

Klann, E., Chen, S.-J., and Sweatt, J.D. (1993) Mechanism of protein kinase C activation during the induction and maintenance of long-term potentiation probed using a selective peptide substrate. Proc. Natl. Acad. Sci. USA 90: 8337-8341

Komori, M. (1993) A novel P450 expressed at the high level in rat brain. Biochem. Biophys. Res. Comm. 196: 721-728

Kozak, M. (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell 44: 283-292

Kozak, M. (1987) An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucl. Acids Res. 15: 8125-8148

Lamb, B.T., Sisodia, S.S., Lawler, A.M., Slunt, H.H., Kitt, C.A., Kearns, W.G., Pearson, P.L., Price, D.L., and Gearhart, J.D. (1993) Introduction and expression of the 400 kilobase precursor amyloid protein gene in transgenic mice. Nature Genetics 5: 22-30

Larson, J., Wong, D., and Lynch, G. (1986) Patterned stimulation at the theta frequency is optimal for the induction of hippocampal long-term potentiation. Brain Res. 368: 347-350

Lathe, R., and Mullins, J.J. (1993) Transgenic animals as models for human disease - report of an EC study group. Transgenic Research 2: 286-299

Lavery, D.J., and Schibler, U. (1993) Circadian transcription of the cholesterol 7α hydroxylase gene may involve the liver-enriched bZIP protein DBP. Genes and Dev. 7: 1871-1884

Li, Y.C., Wang, D.P., and Chiang, J.Y.L. (1990) Regulation of cholesterol 7α-hydroxylase in the liver. J. Biol. Chem. 265: 12012-12019

Li, Y., Erzurumlu, R.S., Chen, C., Jhaveri, S., and Tonegawa, S. (1994) Whisker-related neuronal patterns fail to develop in the trigeminal brainstem nuclei of NMDAR1 knockout mice. Cell 76: 427-437

Liang, P., and Pardee, A.B. (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science 257: 967-971

Linthorst, A.C.E., Flachskamm, C., Holsboer, F., and Reul, J.M.H.M. (1994) Local administration of recombinant human Il-1β in the rat hippocampus increases serotonergic neurotransmission, hypothalamic-pituitary-adrenocortical axis activity, and body temperature. Endocrinology 135: 520-532

Lynch, G., Larson, J., Kelso, S., Barrionuevo, G., and Schottler, F. (1983) Intracellular injections of EGTA block induction of hippocampal long-term potentiation. Nature 305: 20-26

Lynch, G., and Baudry, M. (1987) Brain spectrin, calpain and long-term changes in synaptic efficacy. Brain Res. Bull. 18: 809-815

Lynch, M.A., Errington, M.L., and Bliss, T.V.P. (1989) Nordihydroguaiaretic acid blocks the synaptic component of long-term potentiation and the associated increases in release of glutamate and

arachidonate: an *in vivo* study in the dentate gyrus of the rat. Neuroscience 30: 693- 701

MacMurray, A.J., and Shin, H.S. (1992) The murine MHC encodes a mammalian homolog of bacterial ribosomal protein S13. Mamm. Genome 2: 87-95

Malinow, R., Schulman, H., and Tsien, R.W. (1989) Inhibition of postsynaptic PKC or CaMKII blocks induction but not expression of LTP. Science 245: 862-866

Marshall, R.D. (1972) Glycoproteins. Ann. Rev. Biochem. 41: 673-702

Martini, L., and Melcangi, R.C. (1991) Androgen metabolism in the brain. J. Steroid Biochem. Molec. Biol. 39: 819-828

Mason, I.J. (1994) The ins and outs of fibroblast growth factors. Cell 78: 547-552

Mason, R.P., Shoemaker, W.J., Shajenko, L., Chambers, T.E., and Herbette, L.G. (1992) Evidence for changes in the Alzheimer's disease brain cortical membrane structure mediated by cholesterol. Neurobiology of Aging 13: 413-419

Masters, C.L., Beyreuther, K., Trillet, M., and Christen, Y. (eds) Amyloid protein precursor in development, aging and Alzheimer's disease. (Springer Verlag Berlin)

McInerny, C.J., Kersey, P.J., Creanor, J., Lowndes, N.F., Johnston, L.H., and Fantes, P.A. (1994) Positive and negative regulation of gene expression during the cell cycle by the DSC1^{Sp} complex. Submitted.

McKeown, M. (1992) Alternative mRNA splicing. Ann. Rev. Cell. Biol. 8: 133-155

Miller, W.L. (1988) Molecular biology of steroid hormone synthesis. Endocrine Rev. 9: 295-318

Miller, S.G., and Kennedy, M.B. (1986) Regulation of brain type II Ca^{2+}/calmodulin-dependent protein kinase by autophosphorylation: a Ca^{2+}-triggered molecular switch. Cell 44: 861-870

Monyer, H., Sprengel, R., Schoepfer, R., Herb, A., Higuchi, M., Lomeli, H., Burnashev, N., Sakmann, B., and Seeburg, P.H. (1992) Heteromeric NMDA receptors: Molecular and functional distinction of subtypes. Science 256: 1217-1221

Morgan, E.T., and Gustaffson, J-A. (1987) Sex-specific isozymes of P-450. Steroids 49: 213-245

Mornet, E., Dupont, J., Vitek, A., and White, P.C. (1989) Characterization of two genes encoding human steroid 11β-hydroxylase (P-450$_{11\beta}$). J. Biol. Chem. 264: 20961-20967

Morohashi, K., Sogawa, K., Omura, T., and Fujii-Kuriyama, Y. (1987) Gene structure of human cytochrome P-450 (SCC), cholesterol desmolase. J. Biochem. 101: 879-887

Morris, R.G.M., Anderson, E., Lynch, G.S., and Baudry, M. (1986) Selective impairment of learning and blockade of long-term potentiation by an N-methyl-D-aspartate receptor antagonist, AP5. Nature 319: 774-776

Mountford, P., Zevnik, B., Duwel, A., Nichols, J., Li, M., Dani, C., Robertson, M., Chambers, I., and Smith, A. (1994) Dicistronic targeting constructs: reporters and modifiers of mammalian gene expression. Proc. Natl. Acad. Sci. USA 91: 4303-4307

Mueller, P.R., and Wold, B. (1989) *In vivo* footprinting of a muscle specific enhancer by ligation mediated PCR. Science 246: 780-786

Murthy, K.K., Banville, D., Srikant, C.B., Carrier, F., Holmes, C., Bell, A., and Patel, Y.C. (1990) Structural homology between the rat calreticulin gene product and the *Onchocerca volvulus* antigen Ral-1. Nucl. Acids Res. 18: 4933

Musanti, R., Parati, E., Lamperti, E., and Ghisleei, G. (1993) Decreased cholesterol biosynthesis in fibroblasts from patients with Parkinson's disease. Biochem. Med. Metab. Biol. 49: 133-142

Mutchler, K.J., Klemish, S.W., and Russo, A.F. (1992) A rapid PCR protocol for identification of differentially expressed genes from a cDNA library. PCR Methods and Applications 1: 195-198

Nakamura, T., and Gold, G.H. (1987) A cyclic nucleotide-gated conductance in olfactory receptor cilia. Nature 325: 442-444

Nathans, J., Thomas, D., and Hogness, D.S. (1986) Molecular genetics of human color vision: The genes encoding blue, green, and red pigments. Science 232: 193-202

Nedivi, E., Hevroni, D., Naot, D., Israeli, D., and Citri, Y. (1993) Numerous candidate plasticity-genes revealed by differential cDNA cloning. Nature 363: 718-722

Nelson, D.R., and Strobel, H.W. (1988) On the membrane topology of vertebrate cytochrome P-450 proteins. J. Biol. Chem. 263: 6038-6050

Nelson, D.R., Kamataki, T., Waxman, D.J., Guengerich, F.P., Estabrook, R.W., Feyereisen, R., Gonzalez, F.J., Coon, M.J., Gunsalus, I.C., Gotoh, O., Okuda, K., and Nebert, D.W. (1993) The P450 superfamily: update on new sequences, gene mapping, accession numbers, early trivial names of enzymes, and nomenclature. DNA Cell Biol. 12: 1-51

Nishimoto, M., Noshiro, M., and Okuda, K. (1993) Structure of the gene encoding human liver cholesterol 7α-hydroxylase. Biochim. Biophys. Acta 1172: 147-150

Niwa, O., Kumazaki, T., Tsukiyama, T., Soma, G., Miyajima, N., and Yokoro, K. (1990) A cDNA clone overexpressed and amplified in a mouse teratocarcinoma line. Nucl. Acids Res. 18: 6709

Nojima, H., Kishi, K., and Sokabe, H. (1987) Multiple calmodulin mRNA species are derived from two distinct genes. Mol. Cell. Biol. 7: 1873-80

Noshiro, M., Lakso, M., Kawajiri, K., and Negishi, M. (1988) Rip locus: regulation of female-specific isozyme (I-P-450(16 alpha) of testosterone 16 alpha-hydroxylase in mouse liver, chromosome localisation, and cloning of P-450 cDNA. Biochemistry 27: 6434-6443

Noshiro, M., and Okuda, J.Y. (1990) Molecular cloning and sequence analysis of cDNA encoding human cholesterol 7 alpha-hydroxylase. FEBS Lett. 268: 137-140

Nowack, L., Bregestovski, P., Ascher, P., Herbert, A., and Prochiantz, A. (1984) Magnesium gates glutamate-activated channels in mouse central neurons. Nature 307: 462-465

O'Dell, T.J., Hawkins, R.D. Kandel, E.R., and Arancio, O. (1991a) Tests of the roles of two diffusible substances in long-term potentiation: evidence for nitric oxide as a possible early retrograde messenger. Proc. Natl. Acad. Sci. USA 88: 11285-11289

O'Dell, T.J., Kandel, E.R., and Grant, S.G.N. (1991b) Long-term potentiation in the hippocampus is blocked by tyrosine kinase inhibitors. Nature 353: 558-560

O'Dell, T.J., Huang, P.L., Dawson, T.M., Dinerman, J.L., Snyder, S.H., Kandel, E.R., and Fishman, M.C. (1994) Endothelial NOS and the blockade of LTP by NOS inhibitors in mice lacking neuronal NOS. Science 265: 542-524

Oyler, G.A., Higgins, G.A., Hart, R.A., Battenberg, E., Billingsley, M., Bloom, F.E., and Wilson, M.C. (1989) The identification of a novel synaptosomal-associated protein, SNAP-25, differentially expressed by neuronal subpopulations. J. Cell. Biol. 109: 3039-52

Paine, A.J. (1991) Current status review. The cytochrome P-450 gene superfamily. Int. J. Exp. Path. 72: 349-363

Patterson, P. H. (1994) The role of neuropoietic cytokines in neural development and injury. Abstract H007 in Journal of Cellular Biochemistry. 23rd Annual Keystone Symposia on Molecular and Cellular Biology. Abstract supplement 18B. Wiley-Liss

Paxinos, G., and Watson, C. (eds) (1986) The rat brain in stereotaxic coordinates. 2nd edition Academic Press

Peppelenboch, M.P., Tertoolen, L.G.J., Hage, W.J., and de Laat, S.W. (1993) Epidermal growth factor-induced actin remodeling is regulated by 5-lipoxygenase and cyclooxygenase products. Cell 74: 565-575

Piomelli, D., and Greengard, P. (1990) Lipoxygenase metabolites of arachidonic acid in neuronal transmembrane signaling. Trends Pharmacol. Sci. 11: 367-373

Poulos, T.L., Finzel, B.C., Gunsalus, I.C., Wagner, G.C., and Kraut, J. (1985) The 2.6-Å crystal structure of *Psuedomonas putida* cytochrome P-450. J. Biol. Chem. 260: 16122-16130

Poulos, T.L. (1988) Cytochrome P-450: Molecular architecture, mechanism, and prospects for rational inhibitor design. Pharmaceutical Rev. 5: 67-75

Qian, Z., Gilbert, M.E., Colicos, M.A., Kandel, E.R., and Kuhl, D. (1993) Tissue-plasminogen activator is induced as an immediate-early gene during seizure, kindling and long-term potentiation. Nature 361: 453-457

Ramirez, M.I., Karaoglu, D., Haro, D., Barillas, C., Bashirzadeh, R., and Gil, G. (1994) Cholesterol and bile acids regulate cholesterol 7α-hydroxylase expression at the transcriptional level in culture and transgenic mice. Mol. Cell. Biol. 14: 2809-2821

Rang, H.P., and Dale, M.M. (1987) Pharmacology (Churchill Livingstone, UK) pp177-203

Ridley, A.J., Paterson, H.F., Johnston, C.L., Diekmann, D., and Hall, A. (1992) The small GTP-binding protein rac regulates growth factor-induced membrane ruffling. Cell 70: 401-410

Rogers, J., Ziegler, S., Liu, Q., Breitman, M., Tempest, H., and Ellis, J. (1994) A new, cerebrally expressed gene of the eph/elk receptor tyrosine kinase family. In British Society for Developmental Biology, Spring Meeting Abstract Book.

Rosenmund, C., and Westbrook, G.L. (1993) Calcium-induced actin depolymerisation reduces NMDA channel activity. Neuron 10: 805-814

Sacktor, T.C., Osten, P., Valsamis, H., Jiang, X., Naik, M.U., and Sublette, E. (1993) Persistent activation of the ζ isoform of protein kinase C in the maintenance of long-term potentiation. Proc. Natl. Acad. Sci. USA 90: 8342-8346

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) Molecular cloning: a laboratory manual. (2nd edition) Cold Spring Harbor Laboratory Press.

Sanghera, M.K., Simpson, E.R., McPhaul, M.J., Kozlowski, G., Conley, A.J., and Lephart, E.D. (1991) Immunocytochemical distribution of aromatase cytochrome P450 in the rat brain using peptide-generated polyclonal antibodies. Endocrinology 129: 2834-2844

Sauer, B., and Henderson, N. (1988) Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. Proc. Natl. Acad. Sci USA 85: 5166-5170

Schmid, C.W., and Jelinek, W.R. (1982) The Alu family of dispersed repetitive sequences. Science 216: 1065-1070

Schwartz, J.H. (1993) Cognitive kinases. Proc. Natl. Acad. Sci. USA 90: 8310-8313

Scoville and Milner (1957) Loss of recent memory after bilateral hippocampus lesions. J. Neurol, Neurosurg. Psychiat 20: 11-21

Shimizu, T., and Wolfe, L.S. (1999) Arachidonic acid cascade and signal transduction. J. Neurochem. 55: 1-15

Shivers, B.D., Hilbich, C., Multhaub, G., Salbaum, M., Beyreuther, K., and Seeburg, P.H. (1988) Alzheimer's disease amyloidogenic

glycoprotein: expression pattern in rat brain suggests a role in cell contact. EMBO J. 7: 1365-1370

Schuman, E.M., and Madison, D.V. (1991) A requirement for the intercellular messenger nitric oxide in long-term potentiation. Science 254: 1503-1506

Silva, A.J., Stevens, C.F., Tonegawa, S., and Wang, Y. (1992a) Deficient hippocampal long-term potentiation in α-calcium-calmodulin kinase II mutant mice. Science 257: 201-206

Silva, A.J., Paylor, R., Wehner, J.M., and Tonegawa, S. (1992b) Impaired spatial learning in α-calcium-calmodulin kinase II mutant mice. Science 257: 207-211

Smith, A.G. (1991) Culture and differentiation of embryonic stem cells. J. Tiss. Cult. Meth. 13: 89-94

Sohda, T., Shimizu, M., Kamimura, S., and Okumura, M. (1993) Immunohistochemical demonstration of ethanol-inducible P450 2E1 in rat brain. Alcohol and Alcoholism 1B: 69-75

Somma, M.P., Pisano, C., and Lavia, P. (1991) The housekeeping promoter from the mouse CpG island HTF9 contains multiple protein-binding elements that are functionally redundant. Nucl. Acids Res. 19: 2817-2824

Sommer, B., Keinanen, K., Verdoorn, T.A., Wisden, W., Burnashev, N., Herb, A., Kohler, M., Takagi, T., Sakmann, B., and Seeberg, P.H. (1990) Flip and flop: a cell-specific functional switch in glutamate-operated channels of the CNS. Science 249: 1580-1585

Sonenberg, N. (1994) mRNA translation: influence of the 5' and 3' untranslated regions. Curr. Opinion Gen. Dev. 4: 310-315

Southern, E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98: 503-5

Squire, L.R., and Zola-Morgan, S. (1991) The medial temporal lobe memory system. Science 253: 1380-1386

Squire, L.R. (1992) Memory and the hippocampus: A synthesis of findings with rats, monkeys, and humans. Psychological Review 99: 195-231

Stadler, J., Trockfield, J., Schmalix, W.A., Brill, T., Siewert, J.R., Greim, H., and Doehmer, J. (1994) Inhibition of cytochromes P4501A by nitric oxide. Proc. Natl. Acad. Sci. USA 91: 3559-3563

Stewart, M., and Fox, S.E. (1990) Do septal neurons pace the hippocampal theta rhythm? Trends Neurosci. 13: 163-168

Stewart, C.L., Kaspar, P., Brunet, L.J., Bhatt, H., Gadi, I., Kontgen, F., and Abbondanzo, S.J. (1992) Blastocyst implantation depends on maternal expression of leukaemia inhibitory factor. Nature 359: 76-79

Strittmatter, W.J., Saunders, A.M., Schmechel, D., Pericak-Vance, M., Enghild, J., Salvesen, G.S., Roses, A.D. (1993) Apolipoprotein E: High-avidity binding to β-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer's disease. Proc. Natl. Acad. Sci USA 90: 1977-1981

Stromstedt, M., Warner, M., and Gustafsson, J.A. (1994) Cytochrome P450s of the 4A subfamily in the brain. J. Neurochem. 63: 671-676

Stromstedt, M., Warner, M., Banner, C.D., MacDonald, P.C., and Gustafsson, J.A. Role of brain cytochrome P-450 in regulation of the level of anaesthetic steroids in the brain. Mol. Pharm. 44: 1077-1083

Sung, L.A., Fowler, V.M., Lambert, K., Sussman, M.A., Karr, D., and Chien, S. (1992) Molecular cloning and characterization of human fetal liver tropomodulin. J. Biol. Chem. 267: 2616-2621

Sutcliffe, J.G. (1988) mRNA in the mammalian central nervous system. Ann. Rev. Neurosci. 11: 157-198

Tan, E-C., Leung, T., Manser, E., and Lim, L. (1993) The human active breakpoint cluster region-related gene encodes a brain protein with homology to guanine nucleotide exchange proteins and GTPase-activating proteins. J. Biol. Chem. 268: 27291-27298

Tingley, W.G., Roche, K.W., Thompson, A.K., and Huganir, R.L. (1993) Regulation of NMDA receptor phosphorylation by alternative splicing of the C-terminal domain. Nature 364: 70-73

Upchurch, M., and Wehner, J.M. (1988) Heterosis and resistance to DFP effects on spatial learning in C57xDBA hybrids. Brain Res. Bull. 21: 499-503

Vandaele, S., Nordquist, D.T., Fedderson, R.M., Tretjakoff, I., Peterson, A.C., and Orr, H.T. (1991) Purkinje cell protein-2 regulatory regions and

transgene expression in cerebellar compartments. Genes and Dev. 5: 1136-1148

Van-Hoesen, G.W., Hyman, B.T., and Damasio, A.R. (1991) Entorhinal cortex pathology in Alzheimer's disease. Hippocampus 1: 1-8

Vanier, M.T., Pentchev, P., Rodriguez-Lafrasse, C., and Rousson, R. (1991) Niemann-Pick disease type C : an update. J. Inher. Metab. Dis. 14: 580-595

Vincent, S., Marty, L., and Fort, P. (1993) S26 ribosomal protein RNA: an invariant control for gene regulation experiments in eucaryotic cells and tissues. Nucl. Acids Res. 21: 1498

Volk, B., Hettmannsperger, U., Papp, T., Amelizad, Z., Oesch, F., and Knoth, R. (1991) Mapping of phenytoin-inducible cytochrome P-450 immunoreactivity in the mouse central nervous system. Neuroscience 42: 215-235

Wallace, H., Ledent, C., Vassart, G., Bishop, J.O., and Al Shawi, R. (1991) Specific ablation of non-dividing thyroid follicle cells in adult transgenic mice. Endocrinology 129: 3217-3226

Wang, M.M., Tsai, R.Y.L., Schrader, K.A., and Reed, R.R. (1993) Genes encoding components of the olfactory signal transduction cascade contain a DNA binding site that may direct neuronal expression. Mol. Cell. Biol. 13: 5805-5813

Warner, M., and Gustafsson, J.A. (1994) Effect of ethanol on cytochrome P-450 in the rat brain. Proc. Natl. Acad. Sci USA 91: 1019-1023

Warner, M., Stromstedt, M., Wyss, A., and Gustafsson, J.A. (1993) Regulation of cytochrome P-450 in the central nervous system. J. Steroid. Biochem. Mol. Biol. 47: 191-194

White, K.A., and Marletta, M.A. (1992) Nitric oxide synthase is a cytochrome P-450 type haemoprotein. Biochemistry 31: 6627-6631

Wieraszko, A., Li, G., Kornecki, E., Hogan, M.V., and Ehrlich, Y.H. (1993) Long-term potentiation in the hippocampus induced by platelet-activating factor. Neuron 10: 553-557

Williams, J.H., and Bliss,T.V.P. (1989a) An *in vitro* study of the effect of lipoxygenase and cyclo-oxygenase inhibitors of arachidonic acid on the induction and maintenance of long-term potentiation in the hippocampus. Neurosci. Letts. 107: 301-306

Williams, J.H., Errington, M.L., Lynch, M.A., and Bliss, T.V.P. (1989b) Arachidonic acid induces a long-term activity-dependent enhancement of synaptic transmission in the hippocampus. Nature 341: 739-742

Wilson, R., and Sargent, J.R. (1993) Lipid and fatty acid composition of brain tissue from adrenoleukodystrophy patients. J. Neurochem. 61: 290-297

Winson, J. (1978) Loss of hippocampal theta rhythm results in spatial memory deficit in the rat. Science 201: 160-163

Wisden, W., Parker, E.M., Mahle, C.D., Grisel, D.A., Nowak, H.P., Yocca, F.D., Felder, C.C., Seeburg, P.H., and Voigt, M.M. (1993) Cloning and characterization of the rat 5-HT5B receptor. Evidence that the 5-HT5B receptor couples to a G protein in mammalian cell membranes. FEBS Letters 333: 25-31

Yarden, Y., Escobedo, J.A., Kuang, W-J., Yang-Feng, T.L., Daniel, T.O., Tremble, P.M., Chen, E.Y., Ando, M.E., Harkins, R.N., Francke, U., Fried, V.A., Ullrich, A., and Williams, L.T. (1986) Structure of the receptor for platelet-derived growth factor helps define a family of closely related growth factor receptors. Nature 323: 226-232

Yamagata, K., Andreasson, K.I., Kaufmann, W.E., Barnes, C.A., and Worley, P.F. (1993) Expression of a mitogen-inducible cyclooxygenase in brain neurons: regulation by synaptic activity and glucocorticoids. Neuron 11: 371-386

Zalutsky, R.A., and Nicoll., R.A. (1990) Comparison of two forms of long-term potentiation in single hippocampal neurons. Science 248: 1619-1624

Zaphiropoulos, P.G., and Wood, T. (1993) Identification of the major cytochrome P450s of subfamily 2C that are expressed in brain of female rats and olfactory lobes of ethanol treated male rats. Biochem. Biophys. Res. Comm. 193: 1006-1013

Zhou, M., Small, S.A., Kandel, E.R., and Hawkins, R.D. (1993) Nitric oxide and carbon monoxide produce activity-dependent long-term synaptic enhancement in hippocampus. Science 260: 1946-1950

APPENDIX I CYTOCHROME P-450 EXPRESSION OR ACTIVITY IN BRAIN

P-450	SPECIES	EXPRESSION	METHOD	REFERENCE
CYP1A	rat	increase after partial hepatectomy	Western	Warner *et al.*, 1993
CYP1A1	rat	brain	RT-PCR	Hodgson *et al.*, 1993
CYP1A2	rat	glioma cell line (C6)	RT-PCR	Geng and Strobel, 1993
CYP2A1	rat	glioma cell line (C6)	RT-PCR	Geng and Strobel, 1993
CYP2B1	rat	brain	RT-PCR	Hodgson *et al.*, 1993
	mouse	spinal cord, medulla oblongata, pons, cerebellum, dorsal raphe nuclei, hippocampal pyramidal cells	Immunohistochemistry	Volk *et al.*, 1991
CYP2B2	rat	brain	RT-PCR	Hodgson *et al.*, 1993
CYP2B1/2	rat	induced in brain upon chronic nicotine administration	Western	Anandatheerathavarada *et al.*, 1993
CYP2C6	rat	ethanol-induced olfactory bulb	RT-PCR	Zaphiropolous and Wood, 1993
CYP2C7	rat	glioma cell line (C6)	RT-PCR	Geng and Strobel, 1993
	rat	ethanol-induced brain	Western	Warner *et al.*, 1994
CYP2C11	rat	ethanol-induced brain	Western	Warner *et al.*, 1994

CYP2C11	rat	ethanol-induced olfactory bulb	RT-PCR	Zaphiropolous and Wood, 1993
CYP2C12	rat	ethanol-induced olfactory bulb	RT-PCR	Zaphiropolous and Wood, 1993
CYP2C23	rat	ethanol-induced olfactory bulb	RT-PCR	Zaphiropolous and Wood, 1993
CYP2C	rat	increase after partial hepatectomy and ethanol induction	Western	Warner *et al.*, 1993
CYP2D	rat	brain	RT-PCR	Hodgson *et al.*, 1993
CYP2D4	rat	cerebral cortex	RT-PCR	Komori, 1993
CYP2D6	rat	brain	?	Wolf (pers. comm.)
CYP2E1	rat	glioma cell line (C6)	RT-PCR	Geng and Strobel, 1993
	rat	brain	RT-PCR	Hodgson *et al.*, 1993
	rat	basal ganglia cerebellar cortices ethanol induction: basal ganglia (increase) cerebellar cortices (increase) substantia nigra hippocampus	Immunohistochemistry	Sohda *et al.*, 1993
	rat	ethanol-induced brain	Western	Warner *et al.*, 1994

CYP2E1	rat	neurons and glia neocortex, olfactory bulb, piriform cortex, several thalamic nuclei, cerebellum (exclusively glial) hippocampal pyramidal cells striatum, substantia nigra, brain stem	Immunohistochemistry	Hansson *et al.*, 1990
CYP4A2	rat	cerebral cortex hypothalamic preoptic area cerebellum brainstem	RT-PCR	Stromstedt *et al.*, 1994
CYP4A3	rat	cerebral cortex hypothalamic preoptic area cerebellum brainstem	RT-PCR (can't detect by Northern)	Stromstedt *et al.*, 1994
	rat	ethanol-induced brain	Western	Warner *et al.*, 1994
CYP4A8	rat	cerebral cortex	RT-PCR	Stromstedt *et al.*, 1994
	rat	ethanol-induced brain	Western	Warner *et al.*, 1994
CYP4A	rat	10-fold increase in hypothalamic preoptic area and olfactory lobes during lactation/pregnancy	Western	Warner *et al.*, 1993
	rat	Increase after partial hepatectomy and ethanol induction	Western	Warner *et al.*, 1993
Hct1 (CYP7B?)	rat	hippocampus	*in situ* hybridisation	Stapleton (thesis, 1994)

CYP19	rat	preoptic area medial basal hypothalamus amygdala hippocampus	RNase protection	Abdelgadir *et al.*, 1994
	rat	intense staining in amygdaloid structures and supraoptic nucleus moderate - light staining in paraventricular and arcuate nuclei and hippocampus	Immunocytochemistry	Sanghara *et al.*, 1991

Cytochrome P-450 Activity

P-450 spectral activity detected in human brain: highest activity detected in brain stem and cerebellum; lowest in striatum and hippocampus	Bharmre *et al.*, 1993
Mouse hepatic P-450s (liver microsome membrane preparation) can metabolise brain anandamide (a brain constituent which binds to the cannabinoid receptor)to at least 10 different metabolites	Bornheim *et al.*, 1993
P-450-mediated activity (rat brain microsomal fractions) results in 6-alpha and 7-alpha hydroxylations of the steroid anaesthetic 3α-hydroxy 5α-pregnan-20-one	Stromstedt *et al.*, 1993
Catechol estrogen-forming enzyme activity in brain is inhibited by cytochrome P-450 inhibitors	Paul *et al.*, 1977

This summary is not exhaustive and is possibly biased as in each case, detection of a specific P450 was addressed only with the available probes or antibodies. In addition, some discrepancies were noted in expression detected in different species and in different methods used (*CYP2B1* expression was detected in rat brain by PCR (Hodgson *et al.*, 1993)but not by Western blot (Warner *et al.*, 1994), while a spatially regulated pattern of cyp2b-1 protein was reported in a separate study in mouse brain (Volk *et al.*, 1991).

APPENDIX II CELL TYPE-SPECIFIC INTERACTIONS OF TRANSCRIPTION FACTORS WITH A HOUSEKEEPING PROMOTER *IN VIVO*

The work contained in Appendix II represents that conducted over the first eighteen months of this four year Ph.D. As this has not contributed to the main project presented in the thesis and is not connected to or continued in the main project, it has not been presented in full detail. This work was performed under the supervision of Dr. Patrizia Lavia, a visiting fellow, and Dr. Patrizia Somma from Centro di Genetica Evoluzionistica del CNR, Universita "La Sapienza" (Rome, Italy). The project was discontinued at the Centre for Genome Research on the departure of Dr. Lavia.

Housekeeping genes comprise a subset of genes which have particular regulatory requirements and most likely differ from tissue-specific genes in the mechanisms used to initiate and maintain transcription. Housekeeping genes for the most part are constitutively expressed in all cell types and through all stages of development. Structurally, housekeeping promoters usually lack a canonical TATA box and are invariantly associated with CpG islands. It has been shown that other mechanisms exist for the initiation of transcription from TATA-less promoters; other elements appear to specify the transcription start site (for example, Smale and Baltimore, 1989).

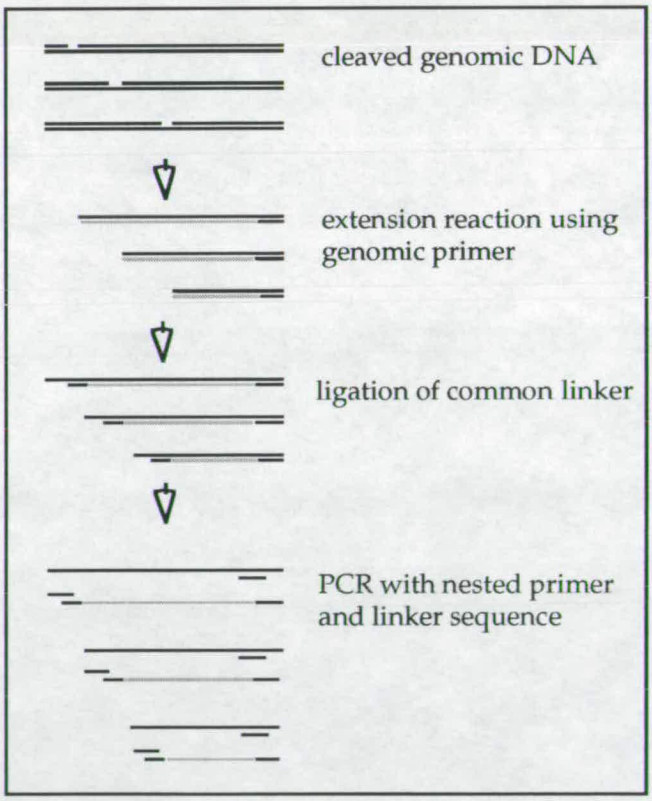
The mouse CpG island, *Htf9*, contains a promoter region responsible for the expression of two divergently transcribed genes which are expressed in all cell types tested (Lavia *et al.*, 1987). In previous studies using *in vitro* protein binding studies and deletion assays, an extensive analysis of this promoter revealed two important points of interest: firstly, *Htf9* contains several potential sites for transcription factor interaction, however only a subset of these are required to drive expression of a reporter gene (Somma *et al.*, 1991a); secondly, there appears to be tissue-specific differential usage of transcription factors by the promoter, depending on the availability of transcription factors supplied by that cell

type (Somma *et al.*, 1991b). Both findings suggested a mechanism of redundancy to ensure transcriptional activation in all cell types.

To test this hypothesis, protein interactions at the *Htf9* promoter were assessed *in vivo* by genomic footprinting, using ligation-mediated PCR (Mueller and Wold, 1989). *In vivo* footprinting permits protein binding sites to be indentified in the context of modulatory factors such as chromatin structure and DNA methylation. In contrast, *in vitro* analyses such as gel retardation, transcription using nuclear extracts and transfection experiments rely on the use of cloned DNA segments. Thus *in vivo* genomic footprinting offers perhaps the most reliable understanding of protein-DNA interactions within a cell. A number of different cell lines of different origins and differentiation states were used to examine the transcription factor occupancy of *Htf9*.

The procedure followed for genomic footprinting is summarised:

1. Cells were exposed to a DNA modifying agent (dimethyl sulphate) which methylates guanine residues except where access is blocked by the presence of a bound protein.
2. Genomic DNA was prepared, followed by piperidine cleavage at every modified base.



A summary of ligation-mediated PCR.

3. Ligation-mediated PCR then requires a DNA polymerase-mediated extension reaction from a primer specific for *Htf9* extending up to each cleavage point to create a double-stranded blunt molecule.
4. At the unknown blunt cleavage point, a double-stranded linker is added by ligation, creating a segment of DNA which can now be amplified by conventional polymerase chain reaction, using a nested genomic primer and a primer based on the linker sequence.
5. The amplified products are fractionated on a sequencing gel, followed by electro-transfer of the DNA to a membrane and probing with the cloned sequence. Footprints are visualised as a disappearance of bands corresponding to occlusion of DMS by template-bound proteins and quite often the appearance of hypersensitive bands at the borders of footprints.

In vivo footprinting confirmed the differential occupancy of the *Htf9* promoter in different cell types. A footprint which overlaps one of the transcription start sites was detected in all cell types except an embryonic stem cell line, cp1, which was subsequently shown not to express the factor (designated GBF). A protection over a CCAAT box was detected only in an hepatic cell line but in none of the others tested. Fibroblasts and a neuroblastoma cell line utilise two high affinity Sp1 sites to maintain transcription, the neuroblastoma cell line showing an additional protection over an AP1 element. A previously undetected interaction was revealed using *in vivo* footprinting, present in all cell types except embryonic stem cells. This occupied site corresponds to an E2F site. Full details are presented in Stapleton *et al.* (1993); appended.

Clearly the maintenance of ubiquitous expression of the *Htf9* transcripts relies on a combination of the available transcription factors supplied by different cellular environments. Questions raised from this work focus on the possible identification of GBF as a novel 'initiator' factor for TATA-less housekeeping promoters and the role of E2F in the regulation of one of the *Htf9* transcripts, *Htf9-a*. Preliminary results from Lavia and colleagues show that *Htf9-a* expression and E2F occupancy of the promoter is associated with cycling cells, while *Htf9-a* is under the negative control of the retinoblastoma gene product in quiescent cells (unpublished).

Nucleic Acids Research, 1993, Vol. 21, No. 10 2465–2471

Cell type-specific interactions of transcription factors with a housekeeping promoter *in vivo*

Genevieve Stapleton, Maria Patrizia Somma[1] and Patrizia Lavia[1]*
AFRC Centre for Genome Research, University of Edinburgh, King's Buildings, West Mains Road, Edinburgh EH9 3JQ, UK and [1]Centro di Genetica Evoluzionistica del CNR, c/o Dipartimento di Genetica e Biologia Molecolare, Università 'La Sapienza', Rome 00185, Italy

Received December 24, 1992; Revised and Accepted April 16, 1993

ABSTRACT

Mammalian housekeeping promoters represent a class of regulatory elements different from those of tissues-specific genes, lacking a TATA box and associated with CG-rich DNA. We have compared the organization of the housekeeping *Htf9* promoter in different cell types by genomic footprinting. The sites of *in vivo* occupancy clearly reflected local combinations of tissue-specific and ubiqitous binding factors. The flexibility of the *Htf9* promoter in acting as the target of cell-specific combinations of factors may ensure ubiquitous expression of the *Htf9*-associated genes.

INTRODUCTION

Over 10.000 genes in higher eukaryotes are thought to encode proteins with 'housekeeping' functions, which are required for survival, growth and duplication of all cells. Such genes are active in all developmental stages and tissues, despite the differences in the transcriptional apparatus of different cell types. Mammalian housekeeping promoters lie within 1–2 kb CG-rich DNA stretches (1–2), exceeding by far the average promoter size. Functional studies indicate that they represent a class of regulatory elements distinct from those of tissue-specific genes. An obvious difference is the absence of the TATA box, which results in heterogeneous transcription initiation (1) often on both DNA strands (3–5). On the other hand, elements that are potential targets of factors with CG-rich recognition sequences (such as Sp1, AP2, MLTF or E2F) occur frequently.

The mouse *Htf9* locus contains a typical housekeeping promoter (3) shared by two genes, *Htf9a* and *Htf9c*, that are transcribed from complementary DNA strands in opposite directions. Both genes are evolutionary conserved, and their expression in all tissues and cell lines suggests that they both encode proteins with basic functions (6). The *Htf9a* gene is the mouse homolog of a novel yeast gene (*SFO1*) recently identified in a search for mutations suppressing the mating deficiency of *fus1* mutants, which are defective in cell fusion (J.Trueheart and J.Thorner, manuscript in preparation). The product of the *Htf9c* gene remains unidentified.

In a previous study the *Htf9* promoter was characterized by combining protein-binding and deletion mapping assays (7, 8). The results indicated two novel features of this promoter. Firstly, we identified multiple factor-binding sites resulting in a complex architecture, yet only a subset was required for expression of a reporter gene in both orientations (7): thus the *Htf9* promoter elements were redundant. This characterization is consistent with results obtained from deletion analysis of unrelated housekeeping promoters (5, 9–12), whose transcriptional activity is also confined to short DNA fragments. Secondly, alternative promoter elements were required for activity in different cell types (8). We have hypothesized that different elements are trans-activated in a cell-specific manner by factors varying from type to type: thus, the apparent redundancy of binding sites might in fact provide the structural basis for ubiquitous expression.

To assess this hypothesis we have examined the organization of the native*Htf9* promoter in different cells by genomic footprinting. The results presented here show that the interactions between factors and single *Htf9* elements in vivo differ in cell lines of different origin and with different levels of specialization. This gives rise to different combinations, involving both ubiquitous and cell-specific factors, which are assembled in a cell type-specific manner.

MATERIALS AND METHODS

Cell lines

The following cell lines were used: mouse C3H/10 T 1/2 (ATCC CCL 226) and NIH/3T3 fibroblasts (ATCC CRL 1658), rat C6 glioma (ATCC CCL 107), mouse S-20Y neuroblastoma (13), rat H4-II-E-C3 hepatoma (ATCC CRL 1600) and human HepG2 hepatocyte carcinoma cells (ATCC HB 8065). Cells were grown in D-MEM medium supplemented with L-glutamine and 10% fetal calf serum under 5% CO_2. Growth conditions for the embryonic stem cell line CP1 and inhibition of differentiation with differentiation inhibitory activity (DIA/LIF) were described in (14).

Gel-shift assays

Protein extracts were prepared following published methods (15, 16). Gel shift assays were carried out as previously described (7, 8) using routinely 1.10^4 cpm of end-labeled probe and 1 μg of double-stranded poly (dI.dC) as a non-specific competitor.

* To whom correspondence should be addressed

In vivo footprinting

The protocol for genomic footprinting using ligation-mediated PCR (LM-PCR) (17) was kindly provided by G. Pfeiffer and generally followed. Cell monolayers were exposed to dimethyl sulfate (DMS) concentrations ranging from 0.05% to 0.5%. In most experiments the effective concentrations fell between 0.1 and 0.3%. Following exposure to DMS cells were collected and DNA extracted by standard methods. After piperidine cleavage the DNA was processed for LM-PCR. To circumvent compressions or stops in regions of high guanine content, Sequenase was replaced with Taq polymerase in the extension step and the temperature was raised to 74°C. PCR-amplified products were run on 60 cm gradient gels and electroblotted onto GeneScreen (Dupont) membranes (18). DNA was UV cross-linked under a 30 Watt lamp from a 80 cm-distance. Single-stranded probes were prepared by primer extension using 4 μg of pH9.2 plasmid containing *Htf9* annealed to 0.5 pmol of the appropriate primer, in the presence of 100 μM dNTPs, 50 μCi of α-^{32}P-dCTP and Sequenase. Filters were hybridized and washed as described in (18). In experiments with whole tissues and DNAse I as the cleaving agent nuclei were isolated through polyamine-free sucrose gradients (19). DNase I digestions and terminal-transferase reactions before LM-PCR were carried out as described in (20).

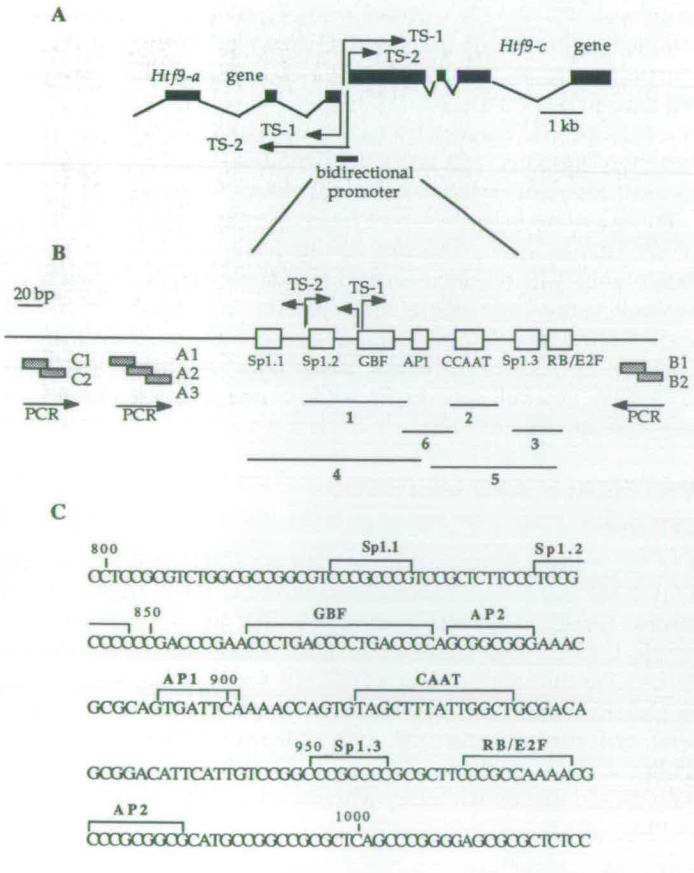


Figure 1. A. Map of the *Htf9* locus, showing the arrangement of the first exons of the *Htf9a* (remaining exons are not located) and *Htf9c* genes. Major transcriptional start sites (TS) of both genes are arrowed. **B.** The bidirectional promoter region in detail: open boxes represent the protein-binding sites identified in vitro (7). The position of the PCR primers (shaded boxes) is shown. Lines represent restriction fragments used in gel shift assays. **C.** Sequence of the *Htf9* promoter. Potential protein-binding sites are indicated.

Oligonucleotides

Two combinations (labeled A and C) of nested primers were copied from the upper strand and one set (B) from the lower strand of *Htf9* (accession number X05830). HPLC-purified primers (Oswel Service, Department of Chemistry, University of Edinburgh) had the following sequences: A1 (5′-ACGGTCT-CGACTAGCTCA-3′), A2 (5′-TAGCTCAGGCGTCGC-TGGC-3′), A3 (5′-TGGCTCGGCGGCCTCTGG-3′), C1 (5′-CCTACCCCACCCGGACCT-3′) and C2 (5′-GGACC-TCGCTGTACCTT-3′), corresponding to positions 734−751, 745−763, 760−777, 680−697 and 692−708; B1 (5′-TGCGCGCGTGGGTCGCCG- 3′) and B2 (5′-TCGCCGGG-AGAGCGCGCT- 3′) correspond to positions 1039−1022 and 1027−1010. The universal linker was composed of a partially complementary 25-mer and 11-mer of respective sequence 5′-GCGGTGACCCGGGAGATCTGAATTC-3′ and 5′-CTA-GACTTAAG-3′. HPLC-purified oligonucleotides for in vitro assays (Genenco Service, Dipartimento di Genetica e Biologia Molecolare, Universitê di Roma) were as follows: 5′-CGTCA-CAAAGAGGCGGGGCTATGCGCATAG-3′ and its complement represent a high-affinity Sp1-binding site; 5′-AGC-TTGAACCCTGACCCCTGACCCCAGCA-3′ and its complement represent the *Htf9* GBF-binding site; 5′-CGCTGT-CGGAGCCAATAAAGCTACAC-3′ and its complement represent the *Htf9* CCAAT site. Double-stranded oligonucleotides carrying the c/EBP (5′-GATCCAGGAATTACGAAATGG-AGGAG-3′) and the E2F- (5′-TAGTTTTCGCGCTTAAA-TTTGA-3′) binding sites were kindly given by A. Vitelli (IRBM Istituto di Ricerche in Biologia Molecolare, Pomezia, Italy) and A. Felsani (CNR, Istituto di Tecnologie Biomediche, Rome, Italy) respectively. AP1- (5′-TTCCGGCTGACTCATC-AAGCG-3′) and AP2- (5′-GATCGAACTGACCGCCCGC-GGCCCGT-3′) binding oligonucleotides were from Promega.

RESULTS

The mouse *Htf9* locus contains two divergent genes that are both ubiquitously transcribed from a shared bidirectional promoter (3). In the promoter region (Fig. 1) several factor-binding sites were identified in vitro (7). Transient expression assays showed that the CCAAT box was a major promoter element in hepatoma cells while being dispensable in fibroblasts; conversely, a major contribution of the Sp1.2 site to the overall promoter activity was observed in fibroblasts (8). To establish wheter these different requirements reflected cell-type specific patterns of trans-activation, we compared the in vivo interactions of nuclear factors with the *Htf9* promoter in different cell lines. To this aim cells were subjected to in vivo footprinting aided by ligation-mediated PCR. Cell lines with different biological characteristics and transcriptional abilities were chosen, though the process of in vitro adaptation may have somewhat quenched the original differences between cell types. The following lines were examined: NIH/3T3 and C3H/10T 1/2 cells, independently derived from mouse embryonic fibroblasts, with the 10T 1/2 cells retaining a broader pluripotency; rat liver hepatoma H4-II-E-C3 cells, expressing a large number of hepatic functions, such as albumin, tyrosine aminotransferase, transferrin, prothrombin; rat C6 cells, derived from a glial tumor and expressing the S-100

protein, typycal of vertebrate neural tissues; mouse S-20Y cells, a cholinergic neuroblastoma line retaining the choline acetyltransferase enzymatic activity (13); mouse CP1 embryonic stem cells, whose differentiation was prevented by differentiation inhibitory activity DIA/LIF (14). DMS titration experiments were initially carried out to establish the effective concentrations in each cell line and for particular sites. Thereafter three concentrations were routinely used in each experiment and a trend analysis enabled us to appreciate the protection of individual sites.

Protected *Htf9* elements in H4-II-E-C3 hepatoma cells

The *in vivo* pattern of the *Htf9* promoter in hepatoma nuclei is shown in Fig. 2. A protection was observed at position 859–875, overlapping one major transcriptional origin of both divergent genes. The protected element consisted of a direct repeat (GGGACTGGGGACTGGGG) recognized by a novel factor called G-binding factor, GBF (8; G. Di Matteo and P. L., unpublished). In vivo the element was protected on both strands and the binding was accompanied by the appearance of an additional band (position 856) at the 5′ edge of the protection (Fig. 2A).

Another footprint was evident over the sequence GAGCCAATAAAGCTA (position 911–925), harboring a CCAAT box (Fig. 2A and 2B). The protection was flanked by hypersensitive bands on either side (positions 940 and 914 respectively, Fig. 2B). Fig. 2C shows the CCAAT site in different cell lines: a distinct footprint was seen in H4 cells only. Because H4 cells are derived from rat hepatoma, rat DNA was sequenced in vivo to assess any divergence which may have occurred between mouse and rat: in the rat sequence one A to G transition (position 926) was found 5′ to, but not affecting, the CCAAT site. Rat C6 cells, derived from a glial tumor, were also analysed as a species-specific control: no CCAAT protection was observed (Fig. 2C). On the other hand, a footprint of similar extension to that seen in H4 cells was mapped in mouse liver nuclei using DNase I (not shown): thus the CCAAT protection is indeed restricted to cells of hepatic origin regardless of the species.

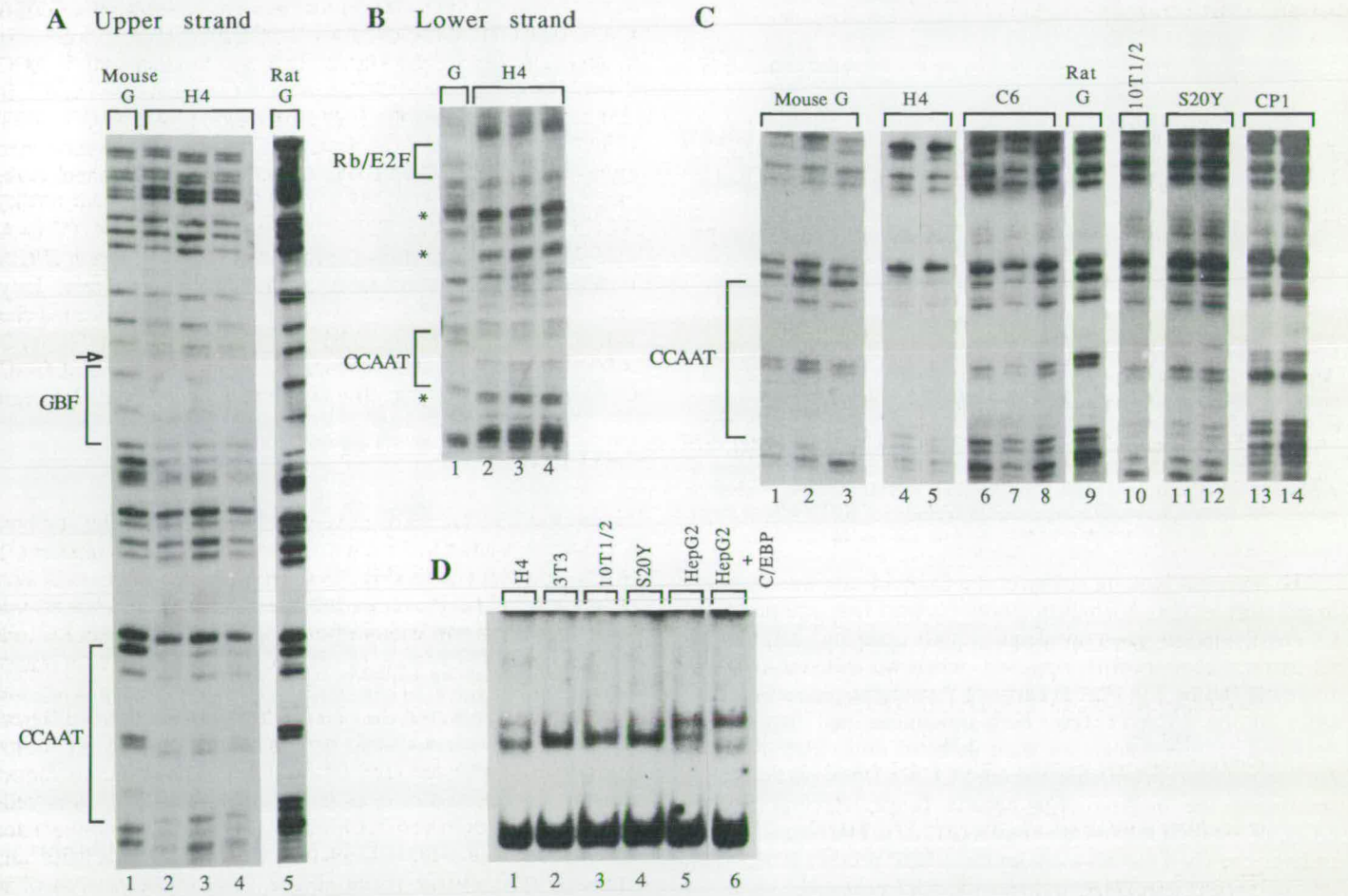


Figure 2. A. Genomic footprints on the *Htf9* promoter in H4 cells. 1. Mouse control DNA; 2–4: H4 cells exposed to 0.1, 0.2 and 0.3% DMS; 5: rat control DNA. Footprints at the GBF and CCAAT sites are bracketed. The arrowhead marks an extraband 5′ to the GBF protection. **B.** Genomic footprint of CCAAT box on the lower strand. 1: mouse control DNA. 2–4: H4 cells treated as for panel A. Asterisks mark DMS-hypersensitive sites. **C.** *In vivo* footprinting of the CCAAT box (upper strand) in all cell types. 1–3: mouse control DNA; 4–5: rat H4 hepatoma cells; 6–8: rat glioma C6 cells; 9: rat control DNA; 10: mouse 10 T1/2 fibroblasts; 11–12: neuroblastoma S-20Y cells; 13–14: mouse CP1 cells. DMS concentrations ranged between 0.1 and 0.3%. **D.** Gel shift assay of *Htf9*- probe carrying the CCAAT box (see Fig. 1). Extract types (4 μg) are indicated above each lane. In lane 6 the reaction was preincubated with a 50-fold excess of c/EBP-binding oligonucleotide prior to addition of the probe.

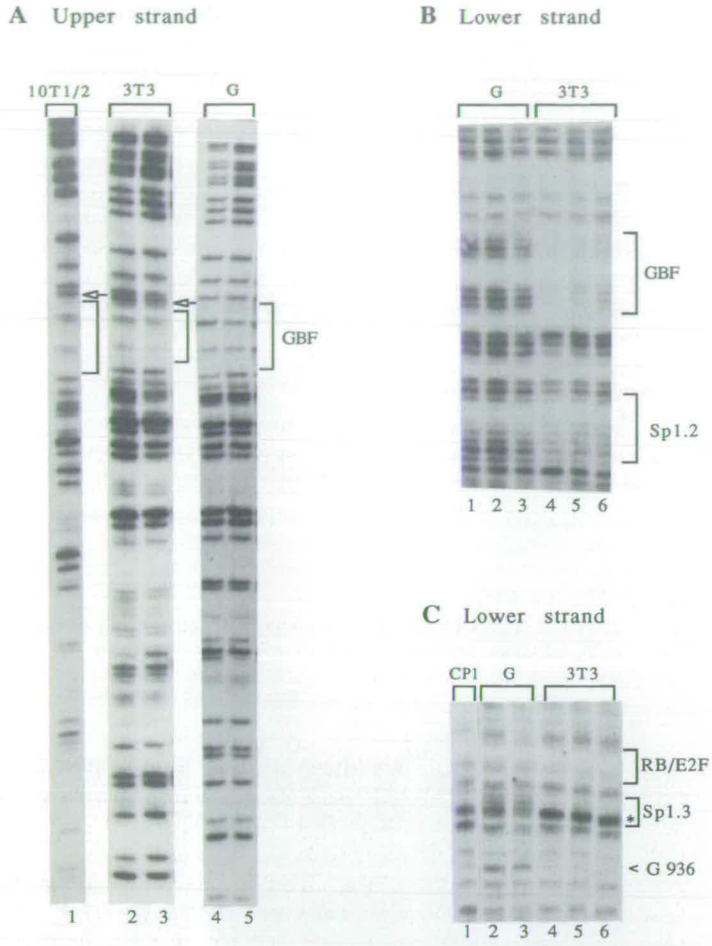


Figure 3. *In vivo* footprints on the *Htf9* promoter in fibroblast lines. **A.** Upper strand. 1: 10T 1/2 cells exposed to 0.3% DMS. 2−3: 3T3 cells exposed to. 0.1 and 0.2% DMS. 4−5: G ladder. The GBF protection is bracketed; the 5′ extraband is arrowed. **B.** Sites Sp1.2 and GBF (lower strand). DMS concentrations ranged between 0.1 and 0.3%. **C.** Sites Sp1.3 and E2F (lower strand). 1. CP1 cells (0.3% DMS); 2−3. G ladder control; 4−6: 3T3 cells treated as for panel B. Asterisks mark hyperreactive bands. Protein-binding sites are indicated. No binding activity was identified corresponding to the G396 protection in 3T3 cells (see text).

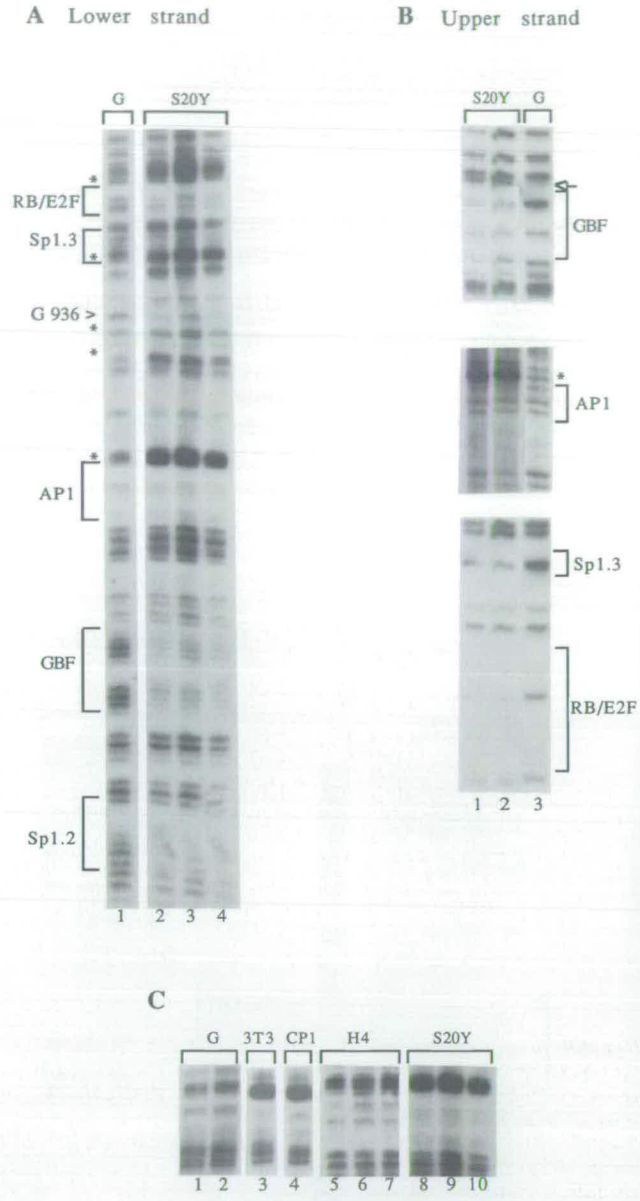


Figure 4. *In vivo* footprints on the *Htf9* promoter in S-20Y neuroblastoma cells. **A.** Lower strand. Lanes 1−2: G ladders from naked DNA controls 3−5: DNA from S-20Y cells exposed to 0.1, 0.2 and 0.3% DMS. Protections are bracketed, hypersensitive bands are marked by an asterisk. **B.** Upper strand. Footprints of the GBF-, AP1-, Sp1.3 and E2F-binding sites. 1−2: S-20Y cells were exposed to 0.1 and 0.3% DMS. 3: G control ladder. The arrowhead marks the extraband 5′ of the GBF protection; asterisks mark hypersensitive bands. **C.** The AP1 site in different cell lines. Cell types are indicated.

The potential binding ability of the CCAAT site was assessed in gel-shift assays. A synthetic 26-mer copied from the protected CCAAT window gave an identical shift upon incubation with all extracts (not shown); however, when we assayed a 40 bp-fragment (probe 2 in Fig. 1) carrying flanking sequences on both sides of the CCAAT box, both ubiquitous and hepatocyte-enriched protein complexes were detected (Fig. 2D): thus the *Htf9* sequences surrounding the core CCAAT box participate in stabilizing the binding of a hepatic factor. Several hepatic CCAAT-binding factors are known (21, 22). The largest family of liver-enriched factors includes the c/EBP protein group (22), whose binding specificity includes CCAAT boxes, the enhancer core motif and c-AMP response elements. We have established that the Htf9-CCAAT factor does not belong to the c/EBP family by site-specific competition assays (Fig. 2D), but have not pursued its identity any further.

Finally, a distinct footprint was seen at positions 964−974 (Fig. 2B) protecting the sequence TTTGGCGG, which matches a high-affinity site for the $p105^{RB}$/E2F complex (23). The protection was also visible on the complementary strand (not shown).

Protected *Htf9* elements in 3T3 and 10 T 1/2 fibroblasts

In both 3T3 and 10T 1/2 fibroblast lines protection of the GBF-binding site was visible on both strands (Fig. 3A and 3B), as in hepatoma cells. Three potential Sp1 sites occur in the *Htf9* promoter (see map in Fig. 1). No interaction was seen with site Sp1.1, whose sequence matches a reported medium-affinity binding site (24). On the other hand, sites Sp1.2 (CTCCGCCCCC, 840−849) and Sp1.3 (CCCGCCCC, 950−957) both match high-affinity sites for Sp1 (24). In vitro binding of fibroblast extracts to both Htf9 sites was very efficient and was competitively inhibited by an excess of Sp1-binding

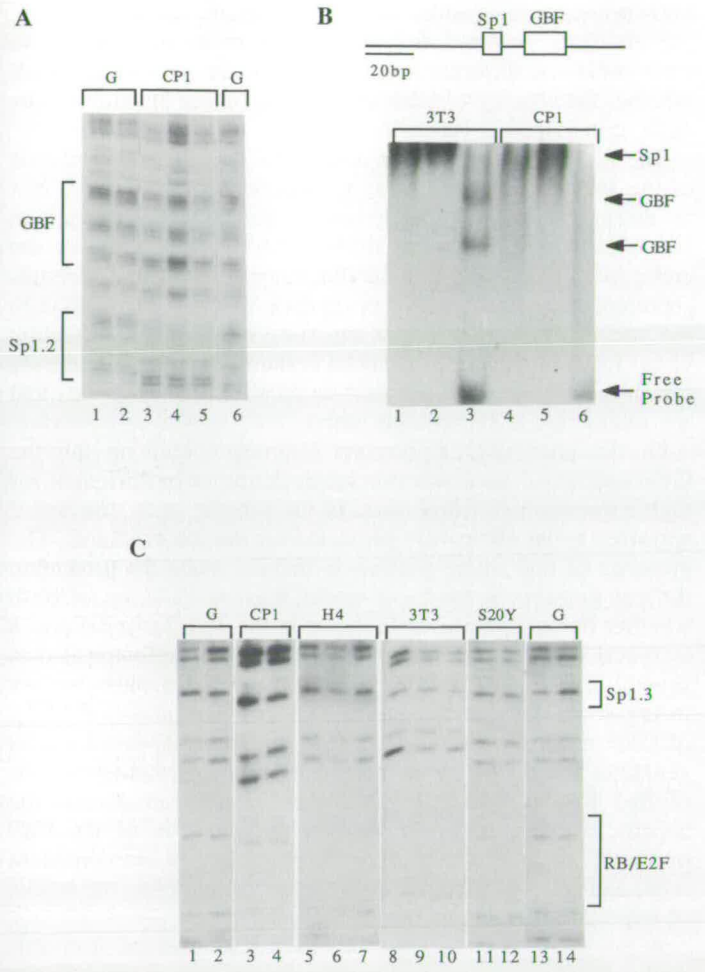


Figure 5. *In vivo* footprints in CP1 embryonic stem cells. **A.** Sp1.2 and GBF sites (lower strand). 1, 2 and 6: G ladder; 3–5: ES cells exposed to 0.1, 0.2 and 0.3% DMS: site Sp1.2 is footprinted while site GBF is not. **B.** Gel shift assay of probe 1, carrying both Sp1.2 and GBF sites. Lanes 1 and 2: 3 μg and 5 μg of 3T3 extract; lane 3: 5 μg of 3T3 extract preincubated with a 50-fold excess of Sp1-binding oligonucleotide: two complexes unaffected by the Sp1 competitor represent different froms of GBF complexes (G.Di Matteo, unpublished); Lanes 4–6: same reactions as in 1–3 using ES cell extracts: no complex is formed after preincubation with the Sp1 oligonucleotide. Migration of the free probe is indicated. **C.** The E2F site in different cell lines (upper strand). Cell types are indicated above each lane. DMS concentrations were between 0.1 and 0.5% DMS. Lanes 1–2 and 13–14: control G ladders. All lines but CP1 show a distinct E2F footprint (bracketed).

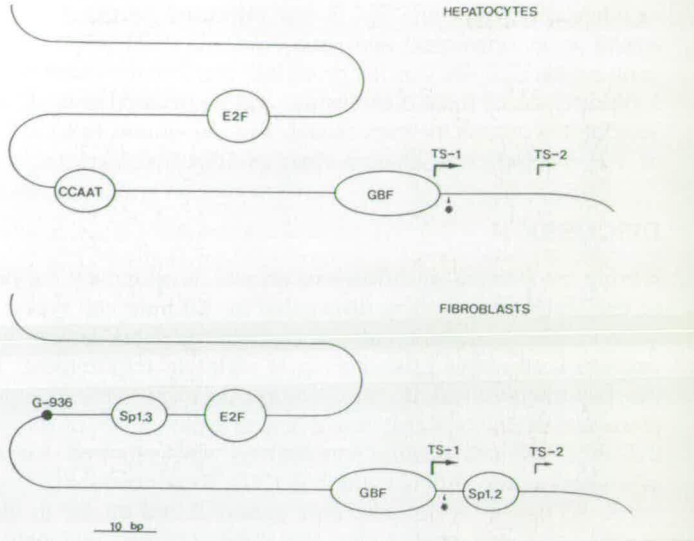


Figure 6. Summary of the *in vivo* protections detected at the *Htf9* promoter in hepatoma cells and in fibroblasts: the orientation of *Htf9-a* transcription is shown. TS-1 and TS-2 indicate the preferred sites of transcription initiation (see ref.3). The asterisk shows the hypersensitive band flanking the GBF protection.

oligonucleotide. In vivo protection of both sites Sp1.2 (Fig. 3B) and Sp1.3 (Fig. 3C) was evident on the G-rich lower strand.

The region surrounding site Sp1.3 showed an altered organization compared to the control ladder on the lower strand. Three footprints were individually distinguishable (Fig.3C). Inspection of the DNA sequence (Fig. 1C) revealed close recognition sites for known factors: site Sp1.3 (950–957) was preceded by the E2F-binding sequence TTTGGCGG (964–974) on the lower strand; the latter overlapped with a potential AP2-binding site (976–985) on the complementary strand. Site-specific competition assays were carried out to disentangle the binding events: binding to Htf9-probe 3 (see Fig.1B) was competitively inhibited by an excess of both Sp1- and E2F-binding oligonucleotides, while remaining unaffected by an AP2-binding competitor (not shown). Thus, both E2F and Sp1 factors are involved in binding to the region and respectively generate the footprints at sites 964–974 (E2F) and 950–957 (Sp1), flanked by hypersensitive bands at positions 946–947. No further binding activity was identified *in vitro*, though the G-residue at position 936 was clearly quenched *in vivo* (Fig. 3C). On the upper strand only the E2F footprint was apparent (see below and Fig. 5C).

Protected *Htf9* elements in S-20Y neuroblastoma cells

In S-20Y neuroblastoma nuclei the GBF-binding site was fully protected on both strands and was bordered by the characteristic 5′ band seen in all other cell lines (Fig. 4A-B). Adjacent to that site an element of similar sequence (TGATTC, position 895–900) to the recognition site for AP1 (TGAGTC) occurs, which was found to be a genuine AP1-binding site by gel-shift assays using a specific competitor (not shown). The AP1 site was protected and flanked by hypersensitive bands in S-20Y cells (Fig. 4A–B). No protection was seen in other examined cell lines (Fig. 4C).

The target sites for Sp1 and E2F showed a similar in vivo occupancy to that observed in fibroblast cells: both Sp1 sites were protected on the lower strand (Fig. 4A); the E2F site was protected on both strands (Fig. 4A–4B); adjacent footprints at sites E2F, Sp1.3 and G 936 on the lower strand generated an in vivo ladder similar to that seen in fibroblasts (compare Fig. 3C and 4A).

Protected *Htf9* elements in CP1 embryonic stem cells

In all somatic cell lines examined thus far the GBF site had been found to be similarly protected in vivo. An exception is represented by the CP1 embryonic stem cell line, in which no GBF footprint was detected (Fig. 5A); the adjacent Sp1.2 site was protected by an independent binding event on the lower strand. In vitro assays confirmed that embryonic cell extracts do not in fact express GBF (Fig. 5B). The pattern of the Sp1.3/E2F region in undifferentiated CP1 cells also differed from that seen

in other cell types: site Sp1.3 was protected on the the lower strand as in fibroblasts and neuroblastoma cells (Fig. 4C); in contrast the E2F site was not protected: therfore the lower strand footprint was of limited extension, and no protection at all was seen on the complementary strand. Fig. 5C shows that the lack of E2F protection is characteristic of ES cells.

DISCUSSION

During the process of differentiation and development the pool of transcription factors is diversified in different cell types. In addition the availability of the so-called general factors also fluctuates, as single promoters have different requirements for the factors involved in basal complexes (25). Housekeeping promoters are ubiquitously active despite these variations. *In vitro* evidence for cell-specific interactions was reported for the promoters of the *Htf9* (8), aldolase (26), B-polymerase (27) and Na,K-ATPase (28) housekeeping genes. Based on the in vitro results, one can predict that the trans-activated elements in housekeeping promoters vary from cell to cell: ubiquitous expression will result from cell-specific patterns of transcriptional activation.

Genomic footprinting provides the most reliable picture of the interactions occurring within the cell (rev. 29) and has proved extremely powerful in pinpointing the onset of activation in inducible (30−32) and developmentally regulated genes (18, 33−35). The technique has also depicted the alternative organization of the X-linked PGK promoter in its active and inactive state (20, 36). In these studies the occupancy of critical cis-active sequences appeared to be an all-or-nothing event related to the acquisition of an open chromatin conformation which triggered transcription. We have employed genomic footprinting to study the native organization of the *Htf9* promoter. We have regarded as informative the *Htf9* sites showing unambiguous protections in at least one cell type—though these may not represent the only functionally relevant elements—and have compared their status in cells expressing different specialized functions. An important point emerging from this study is that *in vivo* protections of the *Htf9* promoter reflected cell type-specific combinations of factors. In contrast, in vitro studies using isolated promoter elements often reflected all potential binding events.

Htf9, like most housekeeping promoters, lacks a TATA box; TATA-less promoters are thought to assemble transcriptional complexes via interactions differing from those occurring at TATA-containing promoters (37) and involving distinct initiators which are only beginning to be identified (38, 39). In this study one major divergent origin of transcription, coinciding with the GBF-binding element, was found to be fully protected on both strands in all differentiated cell lines. The GBF element is included in the shortest DNA fragment required for transcription in both orientations (7). Protection of the GBF site was accompanied by the appearance of a flanking band, which might either indicate a displacement of the G residue 856 at the 5′ border of the protected site, or cleavage of an exposed adenine at position 857, which would also be methylatable by DMS. Both possibilities suggest an altered chromatin structure around the transcription start site. It is noteworthy that the preferrred site of *Ht9-a* RNA initiation is located exactly at 857 (3). It is possible that GBF is a novel member of the protein group involved in transcription initiation in the absence of a TATA box—this question is currently being addressed in our laboratory. Embryonic stem cells, which are the only type lacking GBF, are known to express specific variants of certain 'somatic' factors the elements required for promoter activity often differ in embryonic and differentiated cells. It will be interesting to ask whether the absence of GBF alters transcription initiation of the *Htf9-* transcripts in ES cells.

The *Htf9* promoter carries a CCAAT box in the orientation of the *Htf9-a* gene transcripition. Deletion of the CCAAT box is detrimental for transcription in hepatocytes but not in fibroblasts; in addition the Htf9-CCAAT box can replace the endogenous CCAAT box in the albumin promoter (8). The results reported here show in vivo occupancy of the CCAAT box in the nuclei from rat hepatoma and from mouse liver. Retention of the footprint during liver nuclei isolation—a procedure during which most factors fall off their target sites (19)—suggests that the interaction is remarkably stable.

On the other hand, a promoter fragment containing only the GBF and Sp1.2 sites was previously found to be sufficient for high expression in fibroblasts. In the present study site Sp1.2 appeared to be effectively protected on the lower strand. The presence of one single guanine in the site made the protection difficult to assess on the upper strand, thus we could not establish whether the asymmetrical footprint at the Sp1.2 site reflects a technical or a biological feature. Site Sp1.2 was footprinted in several cell lines (3T3, 10T 1/2, S-20Y and ES cells) but not in hepatoma cells: it is possible that the simultaneous binding of GBF and of a CCAAT-binding factor in hepatoma nuclei generates a steric hindrance incompatible with the further binding of Sp1 to site Sp1.2. Fig. 6 schematically summarizes the hepatocyte-type and fibroblast-type organization of the *Htf9* promoter: together the in vivo footprinting data are consistent with, and retrospectively provide an explanation for, the results of expression assays in mammalian cell types.

A few novel promoter features have also emerged from this study. A cell-type restricted interaction was seen at the AP1 site in S-20Y cells though no functional analysis was carried out in this cell line. The footprint did not require TPA-induction as reported for the in vivo binding of AP1 to the *c-fos* promoter (40). The footprint might be due to a neuronal AP1 subtype indistinguishable from the AP1 factor in gel-shift assays, or to the interaction of neuronal factor(s) with AP1 (or the AP1-related protein) which might stabilize the binding to the *Htf9* promoter.

The distal promoter region showed an interesting organization which has not yet been functionally dissected but may have a significant role in control of *Htf9-a* transcription. A complex pattern was seen in neuroblastoma and fibroblast cells. Alternating hypersensitive and protected elements indicated that more than one protein were involved in the footprint. Sp1 was one of the factors involved and protected the Sp1.3 box on the lower strand in fibroblasts, neuroblasts and embryonic stem cells—the site could not be unambiguously resolved in hepatoma nuclei because of its compression between the E2F and CCAAT protections. The element TTTTGGCGGG was also protected in fibroblasts and neuroblastoma cells. The protected site represents a perfect match to sites selected in vitro by complexes containing the $p105^{RB}$(retinoblastoma) protein (23)—to which the E2F factor is also thought to associate—and is identical to the E2F site in the *c-myb* promoter (rev. 41). The involvement of E2F is compatible with the lack of protection in undifferentiated embryonic stem cells, as published evidence reported very low levels of the embryonal E2F-like factor DRTF1 in teratocarcinoma cells prior to retinoic acid-induced differentiation (42−43). To our knowledge this is the first demonstration o

the in vivo occupancy of an E2F-binding site in proliferating cells. In the light of our observations that *Htf9-a* is expressed in cycling cells in which the $p105^{RB}$ protein was sequestrated by the E1A oncogene, while being repressed in cells transfected with E1A mutants failing to interact with $p105^{RB}$ (A.Bressan, M.Caruso, A.Felsani and P.L., unpublished), occupancy of the $p105^{RB}$/E2F target is very intriguing.

In fibroblast and neuroblastoma nuclei the promoter ladder appeared to be altered beyond site Sp1.3 on the lower strand. No corresponding alteration was seen on the upper strand, which parallels the strand-specificity of the Sp1.3 footprint. The altered organization was apparent in cell types in which both Sp1.3 and E2F-binding sites were occupied, and in our opinion may reflect a local distortion resulting from the simultaneous binding of E2F and Sp1 to adjacent sites, rather than the binding of a novel protein to an independent sequence. Sp1.3 and E2F are separated by the sequence 5′-GGAAGCGCGG-3′, which is part of the recently described SCE element separating the CG-box 1 from the E2F site in the *dhfr* promoter (44). The SCE is not a protein-binding site, yet exerts a repressive function and is thought to generate a distortion of the DNA structure characteristically framed by hypersensitive sites: this would prevent the 'cross-talk' between the Sp1 and E2F proteins when *dhfr* expression is not required. It is tempting to suggest that the altered structure in the distal region of *Htf9* also indicates an element of structural distortion which could serve a similar function and inhibit *Htf9-a* expression in G0 cells.

Our current understanding of transcriptional regulation is based on combinatorial models. The*Htf9* promoter shows an intrinsic flexibility to act as the target of different combinations of factors, which may enable it to direct transcription in cells equipped with different transcriptional machineries. The cooperation between cell-specific and ubiquitous factors in activating different sets of regulatory elements may represent a generalised mechanism maintaining the ubiquitous expression of housekeeping functions during differentiation.

ACKNOWLEDGEMENTS

The *in vivo* footprinting experiments were carried out at the AFRC Centre for Genome Research (Edinburgh, UK) in the laboratory of R.Lathe. We are grateful to G.Pfeiffer and H.Saluz for protocols and help in setting up this work. We also thank A.Smith for ES cells, A.Vitelli and A.Felsani for specific oligonucleotides, A.Bressan and G.Di Matteo for many contributions to the work, and J.Trueheart and J.Thorner for communicating results before publication. R.Perry, H.Saluz and R.Lathe are gratefully acknowledged for their critical comments on this manuscript. During part of this work P.L. was on leave of absence from Centro di Genetica Evoluzionistica del CNR (Rome, Italy) and was partly supported by a Senior fellowship in the framework of the BRIDGE program of the European Communities. This work was supported by Progetto Finalizzato Ingegneria Genetica, Fondazione Istituto Pasteur-Cenci Bolognetti and Programma Bilaterale 104039/04/9210648 Consiglio Nazionale delle Ricerche (Italy), and by the AFRC (UK).

REFERENCES

1. Dynan, W. S. (1986) *Trends Genet.* **6**, 196–198.
2. Bird, A.P. (1987) *Trends Genet.* **3**, 342–347.
3. Lavia, P., Macleod, D. and Bird, A.P. (1987) *EMBO J.* **6**, 2773–2779.
4. Linton, J.P., Yen, J., Selby, E., Chen, Z., Chinsky, J.M., Lin, K., Kellems, R. E. and Crouse, G. F. (1989) *Mol. Cell. Biol.* **9**, 3058–3072.
5. Lennard, A. and Fried, M. (1991) *Mol. Cell. Biol.* **11**, 1281–1294
6. Bressan, A., Somma, M. P., Lewis, J., Santolamazza, C., Copeland, N., Gilbert, D., Jenkins, N. and Lavia, P. (1991) Gene **103**, 201–209.
7. Somma, M. P., Pisano, C. and Lavia, P. (1991) *Nucleic Acids Res.* **19**, 2817–2824.
8. Somma, M. P., Gambino, I. and Lavia, P. (1991) *Nucleic Acids Res.* **19**, 4451–4458.
9. Melton, D., Mc Ewan, C., Mc Kie, A. and Reid, A. (1986) *Cell* **44**, 314–328.
10. Osborne, T., Gil, G.,. Brown, M. S, Kowal, R. and Goldstein, J. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 3614–3618.
11. Atchinson, M.L., Meyuhas, O. and Perry, R.P. (1989) *Mol. Cell. Biol.* **9**, 2067–2074.
12. Smith, M. L., Mitchell, P. J. and Crouse, G. F. (1990) *Mol. Cell. Biol.* **10**, 6003–6012.
13. Kirshner, J. A., Markelonis, G. J., and Max, S. R. (1986) *J. Neurochem.* **46**, 817–821.
14. Smith, A. G. (1991) *J.Tiss. Cul. Meth.***13**, 89–94.
15. Dignam, J.D., Lebovitz, R.M. and Roeder, R. (1983) *Nucleic Acids Res.* **11**, 1475–1489.
16. Andrew, N. and Faller, D. V. (1991) *Nucleic Acids Res.* **19**, 2499.
17. Pfeiffer, G., Steigenwald, S., Mueller, P. R., Wold, B. and Riggs, A. D. (1989) *Science* **246**, 810–813.
18. Saluz, H. and Jost, J. P. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 2602–2606.
19. Mirkovitch, J. and Darnel, J. E. (1991) *Genes Dev.* **5**, 83–93.
20. Pfeiffer, G. and Riggs, A. D. (1991) *Genes Dev.* **5**, 1102–1113.
21. Raymondjean M., Cereghini S. and Yaniv M. 1988, Proc. Natl. Acad. Sci. USA **85**: 757–761.
22. Faisst S. and Meyer S., 1992. Nucl. Acids Res. **20**: 3–26).
23. Chittenden, T., Livingston, D. and Kaemlin, W.G. (1991) *Cell* **65**, 1073–1082.
24. Kadonaga, T., Jones K. A. and Tjian R. (1986) *Trends Biochem.* **11**, 20–23.
25. Parvin, J. D., Timmers, H. T. and Sharp, P. A. (1991) *Cell* **68**, 1135–1144.
26. Concordet, J.P., Maire, P., Kahn, A. and Daegelen, D. (1991) *Nucleic Acids Res.* **19**, 4173–4180.
27. Englander, E. W. and Wilson, S. H. (1990) *Nucleic Acids Res.* **18**, 919–928.
28. Suzuki-Yagawa, Y., Kawakami K. and Nagano K. (1992) *Mol. Cell. Biol.* **12**, 4046–4055.
29. Saluz, H, Wiebauer K. and Wallace, A. (1991) *Trends Genet.* **7**, 207–211.
30. Mueller, P. R., Salser, S. J. and Wold, B. (1988) *Genes Dev.* **2**, 412–427.
31. Saluz, H., Feavers, I. M., Jiricny, J. and Jost, J. P. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 6697–6700.
32. Abravaya, K., Phillips, B. and Morimoto, R. I. (1991) *Mol. Cell. Biol.* **11**, 586–592.
33. Becker, P.B., Ruppert, S. and. Schutz, G. (1987) *Cell* **51**, 435–443.
34. Mueller, P. R. and Wold, B. (1989) *Science* **246**, 780–786.
35. Rigaud, G., Roux, J., Pictet, R. and Grange, T. (1991) *Cell* **67**, 977–986.
36. Pfeiffer, G., Tanguay, R.L., Steigerwald, S. and Riggs, A.D. (1990) *Genes Dev.*4, 1277–1287.
37. Pugh, B.F. and Tjian, R. (1991) *Genes Dev.* **5**, 1935–1945.
38. Smale, T. S. and Baltimore, D. (1989) *Cell* **57**, 103–113.
39. Means, A. L. and Farnham, P. J. (1990) *Mol. Cell. Biol.* **10**, 653–661.
40. Dey, A., Nebert, D. and Ozato K. (1991) *DNA Cell Biol.* **10**, 537–544.44.
41. Nevins, J.R. (1992) *Science* **25**, 424–429.
42. La Thangue, N. B., Thimmappaya, B. and Rigby P. W. J. (1990) *Nucleic Acids Res.* **18**, 2929–2938.
43. Hamel, P. A., Gill, R. M., Phillips, R. A. and Gallie, B. L. (1992) *Mol. Cell. Biol.* **12**, 3431–3438.
44. Pierce, A.J., Jambou, R.C., Jensen D.E. and Azizkhan, J.C. (1992) *Nucleic Acids Res.* **20**, 6583–6587.